Vol. XV, No. 1 *March, 1969*

The Alberta Journal of Educational Research

CONTENTS

The Human Resources Research Council of Alberta: Some Policies, Plans and Perspectives 35
L. W. Downey

Decisions and Curriculum Objectives: A Methodology for Evaluation 17
Thomas O. Maguire

Context and Connotative Meaning in Grade Five 31
Peter O. Evanechko and Robert D. Armstrong

The Effects of Cued Materials and Transformation Variations on Conservation of Number Performance 47
Donald L. Peters and Kenneth Rubin

Ecology of Education 57
Helen L. Pierce

BOOK REVIEWS

O. J. Buros (Ed.), *Reading Tests and Reviews* (Highland Park, New Jersey: Gryphon Press, 1968), Marion D. Jenkinson 64

COMMITTEE ON EDUCATIONAL RESEARCH
The University of Alberta

Editorial Comment

In this issue of the Journal, we present an article by Dr. Lorne Downey who is the Director of the *Alberta Human Resources Research Council.* He states that the purpose of the HRRC is "to inform the process of social policy development." The Council is unique in its multi-disciplinary focus. Its research projects are to be broadly based in several theoretical disciplines, including education.

Maguire discusses various kinds of objectives and their relationship to levels of decision-making. He urges an examination of both implicit and explicit objectives in any curriculum evaluation. He states that mere labelling of the interval between the information used to make a decision and the decision itself will not lead to an understanding of the cognitive sequence involved in the decision-making process. He discusses the validity of mathematical models for the investigation of decision-making. His methodology also includes models for approaching a reconciliation of conflicts in interpretation of objectives and in the relative emphasis on objectives when they are combined to form the basis of instructional units.

Evanechko and Armstrong report an investigation into the "meaning of meaning." The study is based on meaning as a representational mediation process. An instrument using the Semantic Differential was developed to investigate whether or not children's perception of reality is influenced by the context in which a word is examined. Both the method of investigating and the findings make a significant contribution to research in this area.

Peters and Rubin present the results of three studies of conservation of number in young children. One dimension in the study is variation in transformations. The second is variation in the materials to provide cues to highlight the relevant or irrelevant dimensions of the conservation task. They find that variations in transformations seem to have no effect on the subject's performance, but that cued materials seem to have a differential effect over subjects.

Pierce discusses the need to understand the ecology of a community in making educational decisions such as the location and size of schools. She describes how the data were collected and how they might be used in school and community administration.

P.A.L.

Alberta Journal of Educational Research
Vol. XV, No. 1, March, 1969

The Director of the Human Resources Research Council of Alberta describes the origins and background of the Council, the conceptual system on which it is based, the organization and development of its research capability, and its initial research program.

L. W. DOWNEY
Human Resources Research Council of Alberta

The Human Resources Research Council of Alberta: Some Policies, Plans and Perspectives

Underlying Alberta's recently published *White Paper on Human Resources Development* is the implicit assumption that only through the fullest development and conservation of its human resources can a society hope to achieve greatness.

That assumption has imposed upon policy-makers and all the institutions of our society responsibilities and prospects for change such as few of them have ever before contemplated. Indeed, it must now be asked: in light of this newly espoused goal, and in light of on-going events, is it possible that certain social policies and programs may have become obsolete or inadequate? If so, how should these be changed to keep pace with evolving aspirations and conditions? What are the implications of impending policy and program changes for the existing institutions of society? Which of them may be in need of renovation? Which of abolition? And what new institutions may be reqiured?

These are questions of considerable magnitude and far-reaching implications. What is clearly needed to deal with them rationally is a base of knowledge quite in excess of anything now available.

To spearhead and assist in the development of just such a base of knowledge-production activities with knowledge-application activities *raison d'être* is to inform the process of social policy development.

Now that this research organization has been formally launched, certain questions might be posed: What progress has been made to date? What form or structure has the organization taken? What are its research priorities? What research strategies have been adopted? How is H.R.R.C. likely to complement the research capability of the province?

Or, indeed, of the nation? And, finally, how does it propose to link knowledge-production activities with knowledge-application activities in such a way as to make the informing of social policy development a continuous process?

The purpose of this paper is to deal in a very preliminary and tentative way with such questions of policy and procedure. Before I proceed, however, let me provide something of an historical perspective.

Origins and Background of H.R.R.C.

Premier E. C. Manning's *White Paper on Human Resources Development* was tabled in the Alberta Legislature in 1967. In the paper, he declared the intention of his government "to make human resources development a supreme provincial concern" and to create such new institutions as might be required to pursue this intention with vigor.

The same year, the Legislature adopted *An Act to Establish the Alberta Human Resources Research Council*—a quasi-nongovernmental agency whose function it would be "to undertake educational, social, economic and other research relating to and affecting the development and conservation of human resources in Alberta."

By January, 1968, a Council (a governing board representative of government, the academic community, and other sectors of society) had been appointed, as had a director-designate, who was to become the chief executive officer of the Council, effective the following May.

The inaugural meeting of the *Human Resources Research Council* was held in January, 1968; the first major policy meeting was held in June of the same year.

By early autumn a small, interdisciplinary staff of planners had been assembled and the process of translating the Council's goals and objectives into research programs and strategies began.

Though the planning period was too brief to include much by way of pre-planning research, an attempt was made, nevertheless, to impose some rigor and system upon the process. A five-phase (or more accurately, a five-task-area) planning process was conceived and used as a guide to activities.

Phase one was an attempt to lay firm hold of the mandate. This involved a careful analysis of the *White Paper* and the *Act*, an attempt to conceptualize the research domain, and the establishment of general objectives.

Phase two involved the development of a bank of ideas. This task was approached in two ways: first, by gathering (from a large number of individuals and groups) perceptions of emerging or persistent problems in society and suggestions as to how these might be researched; and second, by developing conceptual models which spawned ideas or researchable questions.

Phase three was an attempt to order problems and ideas and to impose various kinds of perspectives and limitations. Included here were studies of research programs under way elsewhere, attempts to identify gaps in knowledge and gaps between knowledge and practice, and efforts to specify aims and priorities.

Phase four was devoted to the description of the various program areas and the specific projects to be included in each.

Phase five, the implementation phase, which is now in progress, involves the design of projects and the mustering of a staff capability (hired, seconded, commissioned or whatever) to complete the various projects.

This planning process resulted, in January 1969, in the publication of a *Prospectus,* which has now become the blueprint for the initial activities of Council.

Development of the Conceptual System[1]

The concept, "human resources research," is itself a complex one. Add to it certain other features of H.R.R.C.'s mandate (to inform social policy, to move research into development, to push development into practice, etc.) and the task of mapping the research terrain assumes staggering proportions.

A series of approaches were adopted and a corresponding series of conceptual frameworks were considered:

The Task Area Approach

When the architects of the human resources legislation attempted to fix the general objects of the research agency, they did so in response to the question: what kinds of research and development tasks ought to be undertaken?

Through this approach, they were able to envision a series of loosely defined task areas: (*a*) the conduct of *fundamental research*, designed to fill gaps in available knowledge; (*b*) the collection and *codification* of available knowledge, in a form useful to policy-makers and practitioners; (*c*) the *development* of plans, materials and inventions, for use by practitioners; (*d*) the initiation and direction of *processes of change;* (*e*) the support and encouragement of *training* for prospective researchers and developers; and (*f*) the *dissemination* of information to all citizens.

Though this view of the role of the research organization is a widely accepted one, indeed, the one that has guided the development of most modern research agencies, it was considered, nevertheless, too diffuse to

[1] Some of the material for this section was taken from Francis S. Chase, "The Human Resources Research Council: Some Perspectives" (mimeo), and A. M. Decore, Organizing for Research: A Systems Approach" (mimeo).

provide the thrust required by a mission-oriented agency, too fragmentary to suggest an acceptable organizational model, and too vague to provide a perspective on the research domain itself.

The Domains of Knowledge Approach

Traditionally, when scholars have attempted to order the world, they have done so in response to the question: what is the nature of the phenomena under study and what are the appropriate strategies of inquiry for the various sets of phenomena? This approach has resulted in the sort of division of research labour which has characterized the university, and had led to the emergence of the various academic disciplines—sociology, psychology, economics, and so on.

Though this view of the research domain is appealing to most academics and researchers, it was not considered appropriate for an organization such as H.R.R.C.—for it does not take account of the interconnectivity of subject matter as it applies to specific human problems. Furthermore, such an approach was thought likely to lead to an undue emphasis upon so-called basic research, to the exclusion of development-inspired and problem or policy-oriented studies—and, in time, to an effort to duplicate the university's research role.

The Social Services Approach

Another approach, the one normally espoused by policy-makers and governments, is to ask: what are the major areas of social service and how might each area be served by research? This approach typically results in the creation of miniature research agencies, each of which is concerned with the problems peculiar to a specific department—health, education, welfare, youth, and so on.

Though such a view of the research responsibilities of H.R.R.C. would appeal to many (particularly policy-makers and practitioners), it is one that inevitably results in yet another kind of fragmentation of effort. Human problems tend to be multi-faceted. A problem which manifests itself in education, for example, is as likely to have its roots in health or poverty as it is to be a purely educational issue.

The Utilitarian Approach

This perspective, the one frequently used to type social policy research organizations, has developed in response to the question: for what specific purposes are various researcheres likely to be undertaken? What are seen as the targets and anticipated outcomes?

Two major types have been identified: (*a*) *policy- and planning-oriented* reseaches, and (*b*) *practice- and development-oriented* researches.

The former refers to research activities to be undertaken for the express purpose of gleaning knowledge to be used as the basis for recommending policy change; the latter refers to studies to be undertaken for the purpose either of filling gaps in the knowledge which is basic to practice or of bridging the gap between the knowledge available and the knowledge used in various fields of endeavor. Recent royal commission studies of education, of bilingualism and biculturalism, and of penology are excellent examples of the former. The research and development activities of the R. and D. Centres and the Regional Laboratories in the U.S.A. are, for the most part, illustrations of the latter.

Though this simple polarization of the mandate of H.R.R.C. also proved useful, and although it permeated much of the subsequent planning, it was soon found to have serious shortcomings. It was suggestive, for example, of a sharp and unreal separation between policy and practice; it implied an overly dependent relationship between the research agency and the process of day-to-day policy-making and practice; and it did not take account of a host of research activities which fall in neither or both divisions.

An Eclectic Approach

The rationale or conceptual framework which finally emerged incorporated some features of all of the above approaches. In this perspective, however, the *forms* of *levels* of research, the anticipated starting points of various research endeavors, rather than the substance, the discipline or the purpose, became the organizing principle.

Four major research forms were envisioned: (*a*) *operations studies* —including a host of preliminary feasibility studies which might be undertaken to inform the process of goal-setting and planning in the research organization itself; (*b*) *base-line studies*—including various kinds of social stock-taking or descriptions, evaluations of existing social programs and institutions, and codifications of accumulating knowledge and emerging practice; (*c*) *research and development activities*—including theoretical, experimental and empirical studies, invention and development activities, and demonstration and field testing; and (*d*) *policy-informing activities*—including traditional dissemination activities and specific policy- and planning-based activities. (These forms are summarized in Figure 1).

It should be noted at this point that this conceptual framework of research forms was developed simply as a guide to planning—not as a device for programming activities or for structuring the organization. In fact, most research programs include many or all of these forms. Some begin with base-line studies and proceed through fundamental research and development to field testing and demonstration. Others begin with

FIGURE 1

HUMAN RESOURCES RESEARCH COUNCIL:

FORM I

OPERATIONS STUDIES

I. *Operations Studies*

The purpose of these studies is to inform the planning activities of the research agency itself. Included here might be such activities as the following:

1. A study of the data-storage and retrieval capabilities of the province and the design of alternative plans for a data system for H.R.R.C.
2. Studies of the research capabilities of the province and the implications for the development of the research capability of H.R.R.C.
3. Studies of social science research organizations, the correlates of research productivity, and the implications for the organization of H.R.R.C.
4. Studies of perceptions of emerging social problems, needed research, and the role of H.R.R.C.
5. Panel and opinion studies (developing the "art of conjecture") to search for possible futures for society and the obstacles to be overcome in attaining them.
6. Others.

FORM II

BASE-LINE STUDIES

II. *Social Stock-Taking*

The purpose of these studies is to obtain descriptions of existing conditions -- to suggest areas of needed research and development and to provide a base against which to assess progress. These may be:

1. Studies of specific *factors* such as opportunity, poverty, etc.
2. Studies of various *areas* or regions such as cities, rural areas, etc.
3. Studies of *groups* such as ethnic, socio-economic, religious, age-level, etc.

III. *Evaluative Studies*

The purpose here is to evaluate various aspects of the social service program. These may take the form of:

1. Evaluations of specific *programs* such as welfare or education.
2. Assessments of specific *institutions* such as the school, the family, etc.
3. Assessments of specific *qualities* in programs or institutions such as the comprehensiveness of a program, the adaptability of an institution, and so on.

IV. *Syntheses and Codifications*

The purpose of these studies is to pull together available knowledge on a particular subject -- either to identify gaps in knowledge (needed research) or gaps between knowledge and practice (needed development work). These may be:

1. Summaries and codification of research data.
2. Reviews of innovative practices.
3. Summaries of emerging theories and informed opinions regarding specific issues.

A RATIONALE FOR A COMPREHENSIVE RESEARCH PROGRAM

FORM III

RESEARCH AND DEVELOPMENT

V. *Fundamental Research*

The purpose of these studies is to add to the store of knowledge on a particular issue. They may take the form of:

1. Theoretical studies.
2. Experimental studies.
3. Empirical studies.

VI. *Invention and Development*

The purpose of these activities is to create inventions or products, based upon available knowledge. Products may take the form of:

1. Materials for use in practice -- for example, instructional materials for education.
2. Processes or procedures for use in such fields as social work, education, etc.

VII. *Demonstration and Field Testing*

The purpose here is to field test a variety of inventions or innovations under local conditions. Inventions may grow out of:

1. Research and/or development activities conducted by H.R.R.C.
2. Reviews of research and emerging theories.
3. Products which grow out of R and D conducted elsewhere.
4. Reviews of trends.

FORM IV

DISSEMINATION ACTIVITIES

VIII. *Dissemination*

The purpose here is to communicate the results of research to other researchers and to policy-makers or practitioners. Communication may take the form of:

1. Newsletters, monographs, research reports, sample materials, etc.
2. Conferences, seminars, clinics, workshops, etc.

IX. *Planning Activities*

The purpose here is to provide specific information to planning agencies, to assist in the interpretation of the information, and to help weigh the probable consequences of alternative courses of action. This may involve close liaison with:

1. An educational planning agency.
2. Social Planning Councils.
3. Other.

theoretical and experimental studies which, in turn, are expected to spawn new developments. Still others are initiated as inventions which are to be field tested and monitored so that they may feed back problems for basic research. And so on.

In each case, the object is to begin at what is considered to be the most promising point—at field testing, if products have already been developed elsewhere; at basic research, if the store of available knowledge is inadequate; at development, if invention and practice appear to lag unduly behind knowledge; and at evaluation or planning, if the needs of a particular client agency are such as to demand immediate evaluation and policy planning.

Organizing and Developing a Research Capability[2]

The organizational alternatives open to an institution such as the *Human Resources Research Council* are innumerable—ranging all the way from the *foundation model* at one extreme to the strictly *in-house model* at the other.

The alternatives must be considered, of course, within the context of the organization's mission. And the more concisely the mission can be stated, the more rational the organizational decisions can become.

In the *foundation model,* virtually all activities are carried on outside the organization. This strategy seems to be most appropriate when (*a*) the mandate is diffuse and changeable, (*b*) the progam emphasis is upon basic research rather than development, and (*c*) the motive is to influence and activate other research agencies, such as universities.

In the *in-house model,* all activities are carried out either by the staff of the organization or under the direction of the staff of the organization. This mode seems to be most appropriate when (*a*) the mission is specific, (*b*) projects involve both research and development, and (*c*) the motive is to influence practice.

Given certain aspects of H.R.R.C.'s mission (the enhancement of the total research capability of the province, the conduct of both basic research and development activities, the informing of social policy and practice, the resolution of the shifting problems of society, and so on), it is difficult to say which alternative, or combination of alternatives, is likely to prove the most viable. The research activities, summarized earlier, are suggestive of a fairly diffuse and shifting, but nevertheless concise, mission; they are suggestive of an intent to influence both patterns of research and patterns of practice; and they are suggestive of a very broad range of research and development functions.

It may prove useful at this point, therefore, to examine another strategy—namely *contracted studies.*

---

[2] Much of the material in this section has been borrowed from R. Gordon McIntosh and John Hudson, "The Human Resources Research Council: Some Organizational Alternatives (mimeo).

The contract or commsision strategy is used by research agencies when (*a*) talents or resources not included in the in-house capability are needed, (*b*) the subject matter of the research is controversial and likely to be "politically dangerous" for the organization, and (*c*) the magnitude of the research task is so great as to be beyond the capability of the inhouse-staff.[3]

Contracted research is to be avoided, however, (*a*) when there are likely to be difficulties of confidentiality, (*b*) when there appear to be dangers of fragmentation or diversion from the mission, and (*c*) when problems of interpretation, supervision, and involvement of clients are to be anticipated.

In addition, there is a danger, in excessive use of the contract mode, that the organization may be distracted from its central mission. Creative researchers tend "to do their own thing"—in their own way, and at their own pace. Hence the problem of anticipating outcomes from contracted research are even greater than those involved in anticipating in-house productivity.

In light of all these considerations, H.R.R.C. has developed a research organization which is structured, in part, on the *foundation model* (inasmuch as it incorporates mechanisms for supporting selected inquiries by independent scholars); in part, as a *contracting agency* (inasmuch as it looks beyond itself for the research talents and expertise that it requires); and, in part, an *in-house R and D centre* (inasmuch as it includes an in-house capability for the conduct of certain kinds of research-development-planning-dissemination activities).

The Initial Research Program

The following is a brief summary of the major program areas and the projects which have been mounted in each. These will continue throughout the 1969-70 fiscal year and, in some cases, thereafter.

PROGRAM AREA 100: OPERATIONS STUDIES

101. A study of *data storage and retrieval* capabilities in Alberta and the design of a data system for H.R.R.C.
102. A study of *research organizations,* the correlates of productivity in such organizations, and the implications for the strategies of H.R.R.C.
103. Preliminary feasibility and codification studies in selected areas (the *inner city, educational facilities,* and others).
104. *Grants-in-aid* of selected studies by scholars in appropriate fields.

PROGRAM AREA 200: OPPORTUNITY

201. A base-line study of inequalities of *opportunity* in Alberta (political, economic, educational, etc.).
202. Second-stage *empirical studies of opportunity* in selected areas or as it applies to specific groups.

---

[3] Our study of Opportunity in Alberta, currently under way, is an excellent example of a contracted study. See Ian Housego and L. Klimek, "Social Opportunity: A Research Program" (mimeo).

PROGRAM AREA 300: EVALUATION STUDIES

Creation of an evaluations task force to develop the tools and technologies of evaluation.

The development of mechanisms for cooperative evaluations of selected *government-sponsored programs* (Welfare, Youth, etc.).

PROGRAM AREA 400: INDIVIDUALITY

401. A base-line study of emerging trends and theories in the *individualization of education* and the extent to which these have influenced Alberta schools.
402. A field testing and second generation development project in *individually prescri ed instruction* (in cooperation with Research for Better Schools, Inc.).
403. Theoretical and experimental studies of the nature of *individuality* and the process of *individuation.*
404. Experimental and developmental activities in *computer-assisted instruction.*

PROGRAM AREA 500: DEMONSTRATION AND FIELD TESTING

501. A small high school project.

Other innovative projects to be determined by field testing advisory committees.

PROGRAM AREA 600: PLANNING

The creation of *mechanisms* (for example, an educational policies committee made up of the Universities Commission, Board of Post-Secondary Education and Department of Education) whereby the results of planning-based studies may be reported directly to the policy-making agency.

PROGRAM AREA 700: DISSEMINATION

In addition to the programmed research projects listed above, two major policy studies are to be initiated: (*a*) a study of Alberta education —an analysis of the relevance to Alberta of the Hall-Dennis Report on education in Ontario; and (*b*) an analysis of social welfare programs and institutions in Alberta. These are to be undertaken in collaboration with appropriate government departments and other agencies.

A Final Word

It is important to realize that H.R.R.C. is a very, very young research organization; it has just completed its first five-month planning period. It is important to realize also that H.R.R.C. is unique among research organizations; there are few precedents or models to quide its development.

It is to be anticipated, therefore, that the goals and functions of H.R.R.C. will shift with shifting conditions in society; that priorities will change with the changing nature of the problems associated with social policy development; and that strategies will evolve in accordance with the working relationships that H.R.R.C. is able to develop with other research agencies in the province.

Finally, it might be noted that two important assumptions, indeed, "articles of faith," guided the creation of the *Human Resources Research Council*:

1. There are, among social problems and programs, interconnectivities which simply must be taken into account in any research and development program having as its goal the improvement of policy and practice. It follows that a broad concept of the research domain, such as *human resources,* is more viable than a series of more narrow concepts, such as *health, education, welfare,* etc., or, indeed, *psychology, sociology, economics,* etc.

2. The nature of policy-based and development-inspired reseach is such that it is not easily accommodated either by a university or by government—though it may well depend upon cooperation with both government and the university for its success. It follows that the concept of a *quasi-nongovernmental agency,* existing on the fringe of both government and the academic community, is more viable than either a university-based or a government-directed social policy research agency.

But these are hypothetical statements. The test of their validity will be our success or failure in developing an integrated, comprehensive research and development program in the field of human resources—and our success or failure in mobilizing the research capability of the province in such a way as to bring a variety of talents and disciplines to the resolution of broadly conceived social problems.

And only time will tell.

Alberta Journal of Educational Research
Vol. XV, No. 1, March, 1969

A methodology for investigating the apparent decision-making behaviour of educators as they deal with educational objectives is described. The importance of investigating this phenomenon in a formative evaluation is stressed and implications for resolving conflicts among various levels of the educational establishment are suggested.

THOMAS O. MAGUIRE
The University of Alberta

Decisions and Curriculum Objectives: A Methodology for Evaluation

Several writers (Smith & Tyler, 1962; Taba, 1962; Dressel & Mayhew, 1954) have discussed the purposes that the statement of educational objectives serve in curriculum design and implementation. Some of the suggested purposes are: to direct ongoing classroom instruction, to guide in the selection of content, to aid in the evaluation of student progress, to direct course evaluation activities, and to aid in the development of new courses. In view of the range of purposes that objectives are called upon to serve, it is not surprising that a good deal of controversy had begun to arise over the form of their statement. On the one side, Mager (1962) has given voice to the demand that the only proper way to state objectives is in terms of explicitly observable behaviours. Taking a less dogmatic view, Atkin (1963) warned that such insistence can lead to the curtailment of writer creativity, and at the extreme pole Eisner (1966) suggested that it is not even possible to state some objectives in behavioural terms.

What appears to be overlooked in the argument is that the purposes served by objectives may dictate the form of their statement. Thus, for the detailed construction of a programmed text, it is necessary to be specific in stating objectives, whereas in the early stages in the development of a curriculum only broadly outlined goals are necessary for providing guidelines to the writers.

Krathwohl (1965) appears to have grappled with the problem of variability in the statement of objectives when he suggested that objectives might be conceptualized as operating at three levels. At level one, the most abstract level, the objectives are stated as broad, general

statements useful for developing programs of instruction or for setting down types of courses that a student should cover. They are general goals toward which several years of education might be aimed. At the second level, the board objectives are broken down into the objectives of units of instruction. These objectives are stated in terms of behaviours and conglomerations of behaviours. At the third and most concrete level, objectives are stated in the specific terms necessary for creating instructional materials. These are the objectives promoted by the programmed instruction group.

Taking a slightly different tack, Taylor and Maguire (1966a) differentiated among levels of objectives in their proposed model for curriculum evaluation. In their model, which was based on a rational-sequential approach to curriculum development, they suggested that the needs of society are interpreted by various social agents (newspaper writers, academicians, government officials, leaders of industry, PTA executives, etc.) into broad educational objectives by curriculum developers. In turn, the behavioural statements were transformed into strategies for the classroom. The students' interaction with these strategies was described as resulting in observable behaviours.

The authors noted that evaluation activities have two components—a measurement component, and a value assessment component. Within the context of the model, the measurement component consisted of establishing the degree of fidelity of each of the translations, i.e., how accurately broad objectives are represented by the behavioural statements, how completely the behavioural objectives are manifested in the strategies, and how congruent the outcomes and objectives are. The judgment role of evaluation with respect to the broad objectives was seen as being concerned with their social worth. It was suggested that the worth of behavioural objectives be judged in terms of their priority to the entire educational program, that the value of strategies be judged in terms of their efficiency and adequacy for bringing about student outcomes, and that judgments of student outcomes be made in terms of their quality and goodness of fit to the objectives.

At closer look, both Krathwohl's ascendency of objectives and the Taylor-Maguire sequential model of evaluation suggest that there are several points at which critical decisions about objectives are made. In analyzing these decisions, and the corresponding decision-makers one might usefully distinguish between the initiators and the recipients of a curriculum package. By initiators one would include the program developers and writers, as well as those social institutions whose pressures bring about the initiation of the curriculum project. The recipients on the other hand, generally (but not always) include teachers and administrators, those agents who do not normally generate the objectives, but who are instrumental in their pursuit.

To the initiators we assign the responsibility of proposing the objectives. Their decisions include resolving questions about content, sequence, relative emphasis to be placed on each of several objectives, kinds of special materials, etc. A second class of decisions would include the strategies necessary for gaining acceptance of the curriculum, and for advising in its implementation. The recipients have a more passive role, and can be thought of as second guessers who make the same kinds of decisions as the initiators, but make them about a program already in existence, and make them with the knowledge of their own school situation. For this group, three sorts of decisions are possible: accept, reject, and modify, although unconditional acceptance is probably rarely found.

It seems almost unnecessary to state that the decisions of the initiators, and in particular the writers, have a profound effect on the curriculum. But the variability in the results among several groups of writers all operating under the same broad guidelines can be enormous. Perhaps the most public example of this is the activities of the Biological Sciences Curriculum Study, (which have been chronicled by Glass, 1964) in which three groups of writers produced sets of materials which had different content, emphasis and sequencing, and yet all of which evolved from the same set of agreed upon issues. Those familiar with the three versions of *High School Biology* will attest to this variability in approach.

Parenthetically, it might be noted that a great service can be done for the technologies of curriculum development and evaluation by those individuals who are intimately involved at the policy making level of curriculum development if they would take the time to write down descriptions of their activities, the kinds of decisions that they make, the kinds of information upon which they base the decisions, the kinds of conflicts that arise within the initiator group and how these conflicts are resolved, etc. These pertinent pieces of information which amount to almost a diary of events that occur in a project director's life would provide invaluable leads to those of us who are concerned with the development of a technology of evaluation.

To return to the BSCS, examples of decisions which recipients make which can affect a new program were also given by Glass. Of course the most basic decision is whether or not to adopt a particular program, but given that this has already been made at a gross level, subsequent issues and decisions can change the tenor of the curriculum. Glass noted that in order to assure optimal results with the laboratory blocks, it was advisable to have classes of twenty-five students or less, and to provide a laboratory assistant. The provision of such conditions is the responsibility of the recipients of the curriculum package and not the initiators, and it may be that failure to allow for small classes of 25 students or

less places the achievement of some of the objectives in jeopardy. Or budget restrictions may demand that priorities of equipment purchase be established in which case failure to purchase a certain piece of apparatus necessitates the elimination of a certain aspect of the course. All of this amounts to the fact that regardless of how well the curriculum is engineered for immediate adoption, certain practicalities of a particular school situation may substantially change the structure and objectives of the curriculum from that which the initiators had originally conceptualized.

Turning to the more explicit decisions made by recipients about objectives, we can see several levels of objectives upon which decisions are focused. The acceptance or rejection of a curriculum can be made most immediately on the basis of its broadest objectives, and it is most likely that it is at this level that the first screening is done by curriculum coordinators, school boards, etc. Given the general acceptance of the curriculum package, more specific decisions such as the relative emphasis to be placed on sets of objectives must be made. Such objectives are the objectives of Krathwohl's level two. To this point the relationship between level of statement of an objective and the level of curricular decision being made seems agreeably monotonic. That is if the recipients are making decisions that will more broadly affect the "appearance" of the curriculum to the students then they are working with more broadly stated objectives. But when we turn to the classroom level, and in particular when we wish to look at the moment to moment activities of the teacher, there is disagreement as to whether the monotonic relation still holds. Some behaviourists would insist that teachers should make a prior statement of objectives at Krathwohl's level three. Such an insistence may be defensible when applied to writers of programmed materials, and perhaps for textbook writers, but it is unreasonable to expect it of classroom teachers. If there is any advantage to the teacher over other methods of presentation, it likely lies in the flexibility of the teacher which allows him to take advantage of the unpredictable occurrences in a classroom discussion and turn them into useful learning experiences. Although it does not necessarily follow that the expression of specific objectives preempts the taking advantage of momentary occurrences, what is known about cognitive set makes such preemption seem probable.

It is more realistic to suppose that the kinds of objectives that the teachers makes most use of in his classroom activities are the objectives of Krathwohl's level two. (This is not to say that all teachers explicitly state their objectives in this form, but it is meant to suggest that the implicit objectives are of this form.) There is no absolute line of demarcation between objectives of levels two and three, but the former tend to be expressed in conglomerations of behaviour or behavioural

constructs such as understanding, appreciating, enjoying, etc., whereas the latter must be stated in terms of observable behaviours. We are not suggesting that behavioural objectives can be discarded from the teacher's purview, for our suspicion is that under certain conditions in which a prescribed method has been found to be the best method of teaching a certain unit of a course, then this method should be followed to the letter by most teachers. Unfortunately, there are rather few such prescribed methods.

For instructional use, Krathwohl's level two objectives are useful for determining the content to be covered in a unit, for selecting the experiences to be arranged for the students, and for guiding in the construction of evaluation instruments. The source of these objectives is usually traceable to the curriculum guides, and the textbooks created by the initiators, although occasionally a creative group of teachers will act as initiators and develop their own. In any case, the teacher has a profound influence on whether or not the objectives of a unit are achieved. This influence may be formal as in the case of the curriculum committee who decides which are the significant aspects of a proposed curriculum. Or, at the classroom level, the teacher may choose to ignore one section of a text and its concomitant objectives. As significant as the formal influence might be, the more important influence is the informal one in which the teacher's attitude toward the importance of an objective exerts an influence on the student. This view is substantiated by Runkel (1956) and Winter (1961, 1965) who have shown that the attitude held by a teacher toward certain objectives of a course are transmitted to the students along with the primary message of instruction, and that these attitudes are reflected in differential student achievement, between students who hold the same attitudes and those who do not.

In short we are contending that both implicit and explicit decisions about the objectives of an educational program determine the shape and direction of the program and are therefore legitimate objects of study for those interested in improving the present state of curriculum evaluation.

A Method for Studying Decision-Making in Curriculum Evaluation

While the study of decision-making as a discipline has moved rapidly ahead since the publication of Von Neumann and Morgenstern's *Theory of Games and Economic Behavior* (1947), rather little fallout has been felt in curriculum evaluation. Possibly the more immediate pressures of developing criteria and methodologies for evaluation have distracted the evaluators' attention from the determination of decision bases. It is important that the rationale for decisions at various levels be made explicit, so that conflicts between initiators and recipients (and between members within those groups) can be resolved, thus maximizing the

chance of achieving the broad objectives that are deemed significant by the society.

A variety of methods for studying decision-making behaviour ranging from simulation to self-report have been used by researchers, but one of the most useful methods was that applied by Hoffman (1960) to the study of clinical decisions. Hoffman suggested that there are two potentially observable elements in a decision, the information upon which the decision is based, and output, or the decision itself. Between these two points is an interval labelled "the decision process." By labelling the interval we are sometimes deluded into thinking that we have some understanding of it. Is the cognitive activity that occurs prior to a decision a process? If so, what are the sequences of cognitive events that occur during the process? What is the duration of the process—when does it start, and when does it end? Our responses to these kinds of questions remind us that "decision process" is a hypothetical construct. It is a label for the interval between information and output, it is not an explanation of it.

The decision process interval can be studied in one of three ways according to Hoffman: by intuition, by physical measurement, (e.g., neurological and biological measurements), and by mathematical models. Hoffman pointed out that the first of these methods is unsatisfactory because of its inherently subjective nature. Further, he noted that while physical measurements may ultimately prove to be fruitful, they are not as yet sufficiently well developed to be of much use. The third method, the use of mathematical models provides the most satisfactory way of studying clinical decision-making. He suggested that we lose nothing by postulating a mathematical analog of the live process and validating the analog by checking for agreement between the input and output of the analog and the same two parameters of the live situation. In this way the fidelity of the model can be determined.

In using the "model method" of studying decision-making behaviour, it is strongly suggested that an iterative cross-validation design be used. The model should be established on one set of data and validation be carried out on an independent set. If the agreement is not high then adjustments are made to the model and the process repeated. Once an acceptable degree of fidelity had been obtained, components of the model can be systematically manipulated to suggest the critical components of the decision processes various classes of decision-makers and to determine how these critical components might differ among them. And in particular the method can be used to explicate differences between initiators and recipients.

Hoffman suggested that the term "paramorphic representation" might be applicable to the postulated model. Just as two substances are said to be paramorphs of each other if their chemical compositions are identi-

cal, but their crystalline structures differ, so too, a model of behaviour is a paramorph of the actual behaviour if it has identical observable qualities (input and output) but different internal structure.

A recent study by the author (Maguire, 1968) illustrates how the paramorphic representation of decision processes might be used in furthering the technology of evaluation. In an attempt to determine how certain classroom decisions about objectives are made, teachers were asked to evaluate a group of heterogeneous objectives on a set of thirty bipolar adjectival scales which had been carefully selected to describe various aspects of the value of educational objectives. After completing these scales, the teachers were asked to make several hypothetical decisions about the objectives. (For example, to indicate the number of classroom hours that should be devoted to the pursuit of an objective, to indicate the degree of commitment that the teacher possessed for the pursuit of the objective, etc.) The values assigned to the objectives were resolved into more basic components and a score for each teacher's judgment of each objective was calculated for each of the value components. Two samples of teachers were used, each of which judged a different set of objectives. Four common value components were found, and in each sample, two additional components were found. The four common components were labelled Subject-Matter Value, Motivational Qualities, Ease of Implementation, and Statement Properties. The first component referred to the objective's value to the discipline. The second referred to the probable reaction from the students, that is, it answers the question, "Will this objective hold student interest?" Ease of Implementation is self explanatory, and Statement Properties referred to the semantic properties of the statement of the objective, that is, is it precise?, is it concrete?, etc.

A search through the literature on interpersonal perception (one field in which models of behaviour had been previously used, and which had characteristics in common with the perception of objectives) revealed that a considerable portion of human judgmental behaviour could be accurately (in the sense described earlier) characterized by the linear model in which decision or perception is "predicted" by the weighted sum of input components. On this basis it was decided to determine how well this model represented the decisions that teachers make about objectives where the input variables are the teachers' perceptions of the values of the objectives. As an estimate of the upper limit of the accuracy of the model, least-squares regression weights and multiple correlations were calculated between the value components and the decisions for each teacher.

Although the accuracy of the model varied over the kinds of decisions made, the median multiple correlations for both samples, for all decisions were greater than .75, which indicated that the linear model

appeared to be an accurate model of this kind of decision behaviour. Of further interest was the finding that while there was a good deal of variation in apparent decision-making behaviour, within a sample of individuals, the behaviour of a given individual was relatively invariant over different decisions. A cursory look at the relationships between biographical variables and decision-making behaviour revealed no generalizable insights.

A further step in the study involved clustering teachers into groups that were homogeneous with respect to apparent decision-making behaviour as described by their derived combinational policies, that is, the paramorphic representations of their decision policies. A by-product of this procedure enabled the researcher to breakdown the complex combinational policies into structurally simpler components, which allowed a more detailed study of what were described as common underlying policies with the specificities of individuals removed. (This was analogous to the factor structure of a group of tests where that part of the variance that is specific to one test has been removed. In the present case the observed individual combinational policies, i.e., the linear models, were resolved into more basic component policies—sort of "factor policies.")

The result of the breakdown into component policies indicated that a surprising similarity existed between those components derived from one set of teachers and objectives, and those derived from the other set. This suggested that combinational policies are not situation specific, that there is a disposition to combine the perceived values of an objective in the same ways regardless of the objective or the individual decision-making situation in which the individual is operating.

The results of the study indicated further that there are several dimensions of value which appear to be salient to the decisions made about objectives. Some, like the value to the discipline, have been acknowledged as important for some time, others such as the motivational qualities of an objective have been generally overlooked. For curriculum evaluation, this finding implies that for a complete evaluation of a particular program, judgments of the various values that an objective possesses must be gathered from pertinent sources. The study of the value bases of decision-making will indicate which of the values are the most important components of decisions for each of the sources.

Although the study described above had the methodological shortcoming of providing no immediate cross-validation for individual teacher combinational policies as was outlined in the strategy for using models, in a broader sense validation was provided by the between-sample consistencies that occurred. The use of built in replication, varying two critical variables (subjects and objectives) provided more inferential strength than is usually put into a correlational study of this sort.

Having established a strategy for studying individual decision-making we need to extend the investigations to the behaviours of writers, school administrators, the public, and even to the students. The major tenet of the Taylor-Maguire model is that interface congruence across all of these agents of curriculum implementation, and in particular between initiators and recipients, is necessary to accomplish the broadest aims of the educational process.

Let us turn now to the causes of interface disagreement and suggest some methods of resolution.

Conflicts in Decisions about Objectives

Let us differentiate between two different kinds of conflicts that will occur in judgments of educational objectives as they are studied using the suggested method. The first kind of conflict will be called perceptual conflict. Two individuals are said to be in perceptual conflict over their judgments of an objective if their conceptions of the meaning of the objective differ. For example, an individual may perceive two objectives of a unit of instruction as being closely related to each other along some meaningful dimension, and as a consequence of this perception may design a program of instruction incorporating this similarity. If another individual does not perceive the connection between the two objectives in the same way, then the two individuals are in perceptual conflict. The result may be that the second individual finds the program of instruction illogical and rejects it.

Distinct from perceptual conflicts are combinational conflicts. The term "combinational conflict" refers to differences in the systems of weights used to combine his perceptions of the various components of the objective in order to reach a decision. In the previous section, various value dimensions were noted. If two individuals perceive an objective as lying in the same position in the space defined by the value dimensions then they are not in perceptual conflict. If, however, they assign different weights to the various value dimensions when making a decision, then they are in combinational conflict.

An illustration of perceptual conflict between groups is provided by the Taylor and Maguire study (1967). They asked subject-matter experts, curriculum writers, and teachers to rate eighteen objectives, gleaned from high school biology texts, on a twenty-seven scale semantic differential instrument. The twenty-seven scales were chosen as representative of the various characteristics of broad objectives. Working with the data from each of the groups separately in the analysis, the investigators calculated the mean scale value for each objective's rating on each of the scales. Correlations between objectives across scales were calculated. After factoring the correlation matrices by the principal axis method, the factors were rotated to the varimax criterion. Goodness

of fit between pairs of factor structures was determined using the Procrustes method outlined by Schönemann (1966). In this way it was possible to determine the congruence of perceptual structures among the three defined groups.

Although there was little difference in closeness of fit between any of the pairs of the three groups, the differences that existed indicated that the experts and writers were most closely attuned to each other and the experts and teachers least closely attuned. If we assume that there is an underlying dimension along which the groups arrange themselves such that adjacent groups have closer perceptual congruence than non-adjacent groups, then the order of the groups would be experts, writers, and teachers. The dimension might be described as "distance from the high school classroom." As well, the order is suggestive of one of the routes along which scientific information is passed from expert to student.

If experts can be considered as representative of one group of individuals potentially capable of applying pressure on the schools in an effort to bring about changes, then they are a segment of the societal press. The order of the groups then follows the ordering of the agents described in the Taylor-Maguire model, and as such, tends to validate the rational-sequential theory of transmission, translation, and transformation of educational objectives postulated by those authors.

As a part of a much larger study, Taylor (1966) had teachers, writers, and experts perform a variety of scaling tasks on various sets of objectives which differed in terms of generality of statement. Using ratio judgments, pair comparisons, and categorical sort methods of scaling, Taylor found that ". . . writers are in a sense mediating between experts' and teachers' viewpoints, although closer themselves to the experts' view [p. 133]." That experts and writers were more congruent than writers and teachers reflects the recent trend in curriculum development movements to draw the writers from the academic world rather than from the school system.

The described studies illustrate the determination of perceptual conflict in a gross way, but finer distinctions are necessary. It is important to specify where the conflicts lie if possible. Again Taylor (1966) illustrates how some finer distinctions might be made. Members of each of his three groups were asked to indicate on a sheet how they would place triplets of objectives so that the relative distances would provide a physical analog of the relationships among the objectives as perceived by the individual. The mean distances in inches obtained from the responses of each group were converted to similarity measures which were factored and rotated. Taylor reasoned that since the dimensions defined by the factors were similarity dimensions (and therefore the space was a similarity space), the loadings of a given

objective on each of the three factors that he found, could be squared and summed to give an indication of the extent to which the objective lay in the common space. That is, for each objective, the square of the distance from the origin of the space to the location of the objective provided an indication of how firmly the objective was implanted in the similarity space. By dividing the square of the distance from the origin by the total variance of the objective the proportion of the variance of the objective lying in the similarity space could be described. Comparing these proportions across teacher, writer, and expert groups, Taylor was able to begin to pinpoint the sources of disagreement among the groups. He found, for example, that teachers and experts differed in their treatments of the objectives "The study of structure verus function" and "The study of the biological roots of behavior," the experts tending to include the former objective in the similarity space and exclude the latter, and the teachers acting oppositely.

A further potentially fruitful method for mapping perceptual conflicts is an adaptation of the semantic differential technique developed by Osgood *et al.* (1957). Those familiar with the technique will recall that Osgood and his associates collected the judgments of a great many individuals on a wide variety of concepts. The judgments consisted of placing the concepts on a series of bipolar adjectival scales. Analysis of the data produced the realization that the affective meaning of the concepts could be adequately described by three broad dimensions: evaluation, potency and activity. Using a few scales representing each of these dimensions, several studies have been carried out which effectively map the affective meaning of concepts. (See for example Jenkins *et al.*, 1961.) When such a procedure is used with individuals, conceptions of a concept differ.[1]

Little use of the semantic differential has been made for the study of broad educational goals although both the Taylor (1966) and Taylor and Maguire (1967) studies have used it for studying science objectives. A recent exception was the study of the value of broad objectives by the author which was cited earlier (Maguire, 1968). In that study, scales were selected to be descriptive of value only, although value was given a rather broad interpretation. The factors that emerged and were replicated were described earlier as Subject-Matter Value, Motivational Qualities, Ease of Implementation, and Statement Properties. Each of these value aspects was defined by a number of related scales, and although it was not a part of Maguire's study, disagreement between individuals' perceptions of the values of objectives could be mapped by reference to their ratings of the objectives on these scales.

Clearly the kinds of dimensions which result from the factor analysis of interscale correlations are dependent upon both the selection of the

[1] See also Evanechko and Armstrong, this issue [Ed.]. (P. 18)

scales and the selection of the concepts used. Thus, in order to delimit the appropriate set of dimensions which might be used as parsimonious descriptions of the connotative meaning of objectives, a broad sample of objectives must be used. Once the relevant (i.e., relevant to curriculum evaluation and to decisions made about curricula) dimensions have been described, we shall be in a position to determine whether or not two disagreeing individuals are in fact considering the same objective or not. In a sense we are saying that by mapping the semantic structure of objectives, we should be able to increase the fidelity of communication between the initiators and recipients of the curriculum package and that this mapping task falls into the area of responsibility of evaluators, whose lot it is to determine the points of success and failure of educational programs. Thus, if a curriculum fails because the teacher's perceptions of the meaning of an objective are different from the writer's, it is the evaluator who will suggest that the teachers may require a short preliminary course outlining the writer's intentions.

Up to now almost all studies having to do with conflicts between significant agents in curriculum development have been confined to studies of perceptual conflict. If both parties act rationally under conditions of perceptual conflict, then the disagreement should disappear once the point of miscommunication has been clarified. This hopeful thought with its suggestion of immediate resolution of conflict may account for the greater amount of attention that has been paid to perceptual conflicts over that paid to combinational conflicts, for the assumption of rationality in the latter case seems tenuous.

Despite the difficulties, we must turn to conflicts of combinational policies. We have seen how combinational policies have been paramorphically represented in clinical decision studies (Hoffman, 1960) and in decisions about objectives (Maguire, 1968). What we do not have is evidence concerning the use of the results of such methods for the resolution of conflicts. While we need not hesitate to use models for predicting decision-making behaviour, we must be prepared for disappointment when using them in conflict resolution. With this caution in mind, let us take a closer look at some of the implication for curriculum development of combinational conflicts.

For illustration, the previously described study conducted by Maguire (1968) in which the perceived values of objectives were said to be combined by teachers into decisions about the objectives seems useful. Situations arose in which a group of teachers were in substantial agreement on the kind of value that an objective possessed—for example, an objective was rated high on subject-matter value, but low on ease of implementation. (The members of the group were in perceptual agreement.) The results of the study showed, however, that one subgroup appeared to weight highly the subject-matter value and made decisions

favourable to the implementation of the objective. In a sense the members of this group were saying "I know that this objective will be difficult to implement, but its achievement is so central to the subject that it must be pursued." The other subgroup tended to place more emphasis on the ease of implementation value, and seemed to say, "Although this objective is an important part of the subject, the time and effort that would be necessary for its implementation could better be spent on other areas." This, then is an example of combinational conflict.

The question that emerges is how can the evaluator use the kind of information emanating from the paramorphic representation of decisions to increase the chances of the student being able to achieve the goals set by the society. One direction which an evaluator might take in the above example of conflict would be to encourage the curriculum developers to develop and communicate methods of implementation so that the plot of the objective on the ease of implementation dimension would move from difficult to easy. Thus, even though the weighting systems remained unchanged (and in a sense are still in conflict), the end result would be congruent decisions. The first subgroup would implement the objective because of its importance to the subject-matter area, and the second group would implement it both because of its subject-matter significance *and* because of the ease of implementation.

The example makes the resolution of combinational conflicts appear deceptively simple. What may happen in fact is that the adjustments prescribed from the analysis produce no change in the decision made; that the paramorphic representation of decision processes is not a sufficient basis from which workable adjustments of the curriculum can be made. Further investigation by evaluators is needed here.

Summary

In summary, an attempt has been made to outline one of the directions which is being taken by evaluators in their investigation of decision-making in curriculum development and implementation. In addition future directions of needed research have been indicated. A plea has been made to urge evaluators to make their experiences public in a manner reminiscent of clinical methodologists so that all might benefit. Finally, a call has been made for behavioural scientists to become more involved with the study of curriculum evaluation as a technology with the hope that new techniques, instruments and strategies will result from this increased involvement.

References

Atkin, J. M. Some evaluation problems in a course content improvement project. *Journal of Research Science Teaching,* 1963, 1, 129-132.

Dressel, P. R., & Mayhew, L. B. *General education: Exploration in evaluation.* Washington: American Council of Education, 1954.

Eisner, E. W. Educational objectives: Help or hindrance. Paper presented at the 50th annual meeting of the American Educational Research Association, Chicago, February, 1966.

Glass, B. Renascent biology: A report on the AIBS Biological Sciences Curriculum Study. Chapter 7 in R. W. Health (Ed.) *New curricula.* New York: Harper Row, 1964.

Jenkins, J. J., Russell, W. A., & Suci, G. J. An atlas of semantic profiles for the 360 words. *American Journal of Psychology,* 1958, 71, 688-69.

Krathwohl, R. Stating objectives appropriately for program, for curriculum and for instructional materials development. *Journal of Teacher Education,* 1956, 16, 83-92.

Mager, R. R. *Preparing objectives for programmed instruction.* San Francisco: Fearon Publishers, 1962.

Maguire, T. O. Value components of teachers' judgments of educational objectives. *AV Communcation Review,* 1968, 16, 63-86.

Hoffman, P. J. The paramorphic representation of clinical judgment. *Psychological Bulletin,* 1960, 57, 116-131.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. *The measurement of meaning.* Urbana, I11.: University of Illinois Press, 1957.

Runkel, P. S. Cognitive similarity in facilitating communications. *Sociometry,* 1956, 19, 178-191.

Schonemann, P. H. A generalized solution of the orthogonal procrustes problem. *Psychometrika,* 1966, 31, 1-10.

Smith, E. R., Tyler, R. W., & the Evaluation Staff. *Appraising and recording student progress.* New York: Harper & Bros., 1942.

Taba, H. *Curriculum development: Theory and practice.* New York: Harcourt, Brace & World, 1962.

Taylor, P. A. The mapping of concepts. Unpublished doctoral dissertation, University of Illinois, 1966.

Taylor, P. A., & Maguire, T. O. A theoretical evaluation model. *Manitoba Journal of Educational Research,* 1966, 1(2), 12-17.

Taylor, P. A., & Maguire, T. O. Perceptions of some objectives for a science curriculum. *Science Education,* 1967, 51, 489-493.

Von Neumann, J., & Morgenstern, O. *Theory of games and economic behavior.* Princeton, New Jersey: Princeton University Press, 1947.

Winter, W. D. Student values and grades in general psychology. *Journal of Educational Research,* 1965, 55, 331-333.

Winter, W. D. Values and achievement in a freshman psychology course. *Journal of Educational Research,* 1961, 54, 183-186.

Alberta Journal of Educational Research
Vol. XV, No. 1, March, 1969

This study investigated the influence of two possible contexts, social studies and literature, on the connotative meanings of concepts. It was concluded that context does influence the meanings of concepts and provides individuals with differing perceptions of reality. Implications for instruction are suggested.

PETER O. EVANECHKO
and
ROBERT D. ARMSTRONG
The University of Alberta

Context and Connotative Meaning in Grade Five

Philosophers, linguists, psychologists, and educators have over the centuries sought to determine the meaning of meaning and the role of meaning and language in perception, behaviour, and culture since ". . . Human existence is welded to language" (Lotz in Saporta, 1961, p. 1). Educators, in general, have been concerned with the part language plays in how and what the student learns while teachers of the language arts, in particular, have sought to make language learning and vocabulary development as meaningful as possible.

The hypothesis of language relativity has, in this instance, a special significance for, if reality is a function of meaning, then variations in meaning due to situations or contexts is of particular concern to all educators. Whorf (in Saporta, 1961, p. 465), in discussing the concept of linguistic relativity, concludes that each language embodies a particular world view. The speakers of a language are partners to an agreement to see and think of the world in a certain way—not the only possible way. Thus, different languages or different meanings within languages structure different realities. The complexity of contemporary education and society demands specialization in language function. This specialization results in words taking on particular meanings in different contexts. The nature and effect of these special meanings is of particular concern to the teacher.

The Problem

The study of the meaning of words is concerned with establishing the relationship of the words as signs or symbols to things, situations, and behaviours, i.e., to show the relationship of signs to that which they stand for, their significates. The study of word meaning is thus concerned with the role of the organism in mediating the relation between the sign and the significate. According to the view of language as a mediational process, "words represent things because they produce in human organisms some replica of the actual behavior towards these things, as a mediation process" (Osgood, 1957, p. 7). Words represent reality by virtue of the manner in which their meanings were learned. The contexts of the individual's learning history influence the representing property or meaning of the word, making meaning a function of the situational context and the syntax of the academic discipline within which language is used. Affective or connotative meaning underlies denotative meaning attached to concepts in language and, therefore, structures the individual's perceptions of the referents of these concepts. Any variation in word meaning for individuals will structure reality differently for them and is, therefore, an area of significant concern for educators.

This study was designed to investigate the influence of two possible contexts, social studies and the literary context of the reading program, upon the meaning of concepts and, consequently, upon the perception of reality. The study was largely original in its theoretical assumptions and was exploratory and descriptive in nature. Not all variables were adequately controlled: variations in syntactic complexity, in particular, may have biased the results.

Design of the study

Theoretical Position

In this study, meaning was considered to be a representational mediation process wherein words represent things through the production in humans of some part of the actual behaviour towards things for which they are symbols. The meaning of words and concepts becomes developed through the contiguous relationship of sign and significate or of assign and significate where, in the latter case, a word is assigned meaning through a process of contiguous relation with other words. Situational factors comprising the totality of conditions which influence a behavioural event were considered to form a contextual screen operative in the selection of differential stimuli for the organization of meaning. The meaning of a word accordingly varies with the context in which it occurs. Figure 1 illustrates the putative relationship between meaning and context.

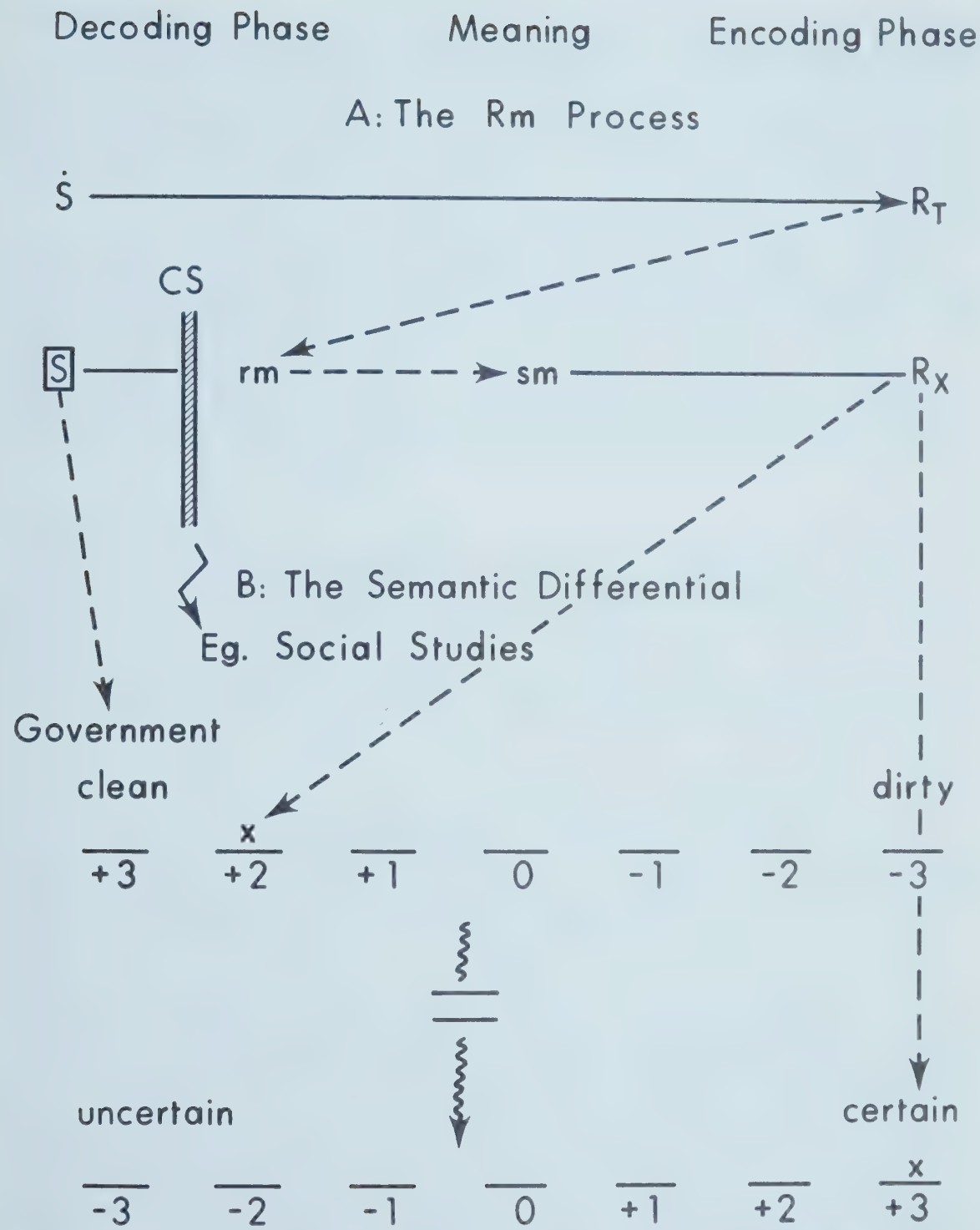


Fig. 1. Diagrammatic representation of the theoretical relationship between the Rm process and the operation of the semantic differential when social studies context is the determining variable. (After Osgood [1957]).

The stimulus producing process (rm - - - - - sm) generated during the decoding phase is representational because it is part of the same behaviour (R_T) produced by the significate (S) and is thus its symbolic semantic property. It is mediational because the self-stimulation (sm) (the encoding of the response) produced by making this short-circuited reaction can now become associated with any instrumental acts (Rx)

which "take account" of the thing signified resulting in suitable responses being evoked. The nature of the response is predicated upon the meaning of the sign which in turn is determined by those portions of the response to the significate passed through the contextual screen (CS). The contextual screen will vary with the situation within which language is used producing different signification for words.

Osgood (1957) categorizes this kind of meaning as connotative meaning. It is considered that state of affairs in a sign using organism that is a necessary subsequent condition (rm) to the decoding of signs and a necessary antecedent condition (sm) to the encoding of signs. This comprises the defining qualities of the category or class named by the word and the affective responses aroused in the minds of its users.

The Sample

To obtain a sample representative of the small town and rural school population of central Alberta within the grade level selected for this study, all grade five students registered in the schools of Leduc County, Alberta, Canada were employed. This group comprised approximately three hundred seventy individuals in seventeen classrooms from nine elementary schools.

The Test Instrument

Osgood (1957) has designed a psychometric method, the Semantic Differential, for providing an index of the linguistic responses or the representational mediation process. Figure 1B illustrates the application of this instrument to the measurement of meaning in different contexts. The context is established by defining and limiting for the subject the nature and scope of the totality of conditions under which he is to respond. In this study, the contexts established were those of two subject areas in elementary school, viz. social studies and literature.

When rating a concept on a given scale, the subject is believed to fix his connotative meaning of the concept in semantic space. By assigning the concept a point on the scale, he establishes both a direction from the origin, identified with a quality of the meaning and a distance from the origin, identified with the intensity of the meaning.

The instrument proper consisted of two forms, one for each context. Each form comprised twenty scales using a seven point bipolar continuum. Adjective scales representing the ways in which meaning may vary in these contexts were selected. The ten stimulus words were identical in both contexts and were selected on the basis of familiarity to the subjects and significance in the area being investigated. The ten stimulus words were government, land, power, wrong, command, money, honor, kill, peace, and yourself.

Procedure

A single treatment and one form of Semantic Differential were administered to each of the seventeen classroom groups in the study according to a randomly assigned order. The treatment consisted of establishing the appropriate situational context in terms of the subject area, i.e., a social studies context and a literature context. This was achieved by the teacher's inclusion of the treatment within the regular activities in the two subject areas. Furthermore, sentences selected at random from written material (textbooks) in each area served to set the context by introducing the stimulus word.

A control group consisting of six classrooms was also identified. This group was administered both forms of the test at an interval of two weeks. Treatment for each of the two groups within the control group was also randomly assigned.

Treatment of the Data

Subject reliability coefficients were obtained for students in the control group. These provided statistical evidence of the effect of treatment upon responses in a test-retest situation.

Factor analysis was obtained by the principal axes method. To obtain simple structure (Thurstone, 1947) the factors for each form of the test were rotated according to the Varimax Criterion (Kaiser, 1958). The rotated axes thus obtained constituted the data for the factorial structures which were interpreted and compared. There are no tests of significance for factors in factor analysis, consequently the Ahmavaara Factor Match Test was employed to test congruence between factor structures and thus determine influence of treatment.

Findings

Control Group Findings

The reliability of responses of individual subjects was obtained by computing the correlations between scale ratings for each concept on one form of the instrument with those on the second form. It was hypothesized that the effect of experimental treatment would cause individuals to respond differently to the same concepts resulting in a lower correlation between the two responses than in the case of a test-retest using the same context. The mean r subject reliability value for the control group was .431. Studies conducted by Osgood (1957), Lilly (1965), Norman (1959), and Jenkins, Russell, and Suci (1958) found subject reliabilities "very high" ranging from .70 to .99. Thus in Lilly's words, "In the absence of experimental treatments . . . ratings . . . remain quite stable over time" (Lilly, 1965, p. 31). It was concluded that the effect of the treatment had been to decrease the coefficient of correlation between responses, providing statistical evidence for the effects of context upon word meaning.

The Ahmavaara Factor Match Test produced cosines of angles between axes of the orthogonal dimensions of meaning ranging from .8217 to .9794. In agreement with other researchers, Lilly (1965) and Eastman (1965), it was concluded that while there was a basic similarity in meaning between contexts, there were also measurable differences. Operating on this conclusion, the differences in factor structures of meanings for the experimental groups were assumed to be real.

Experimental Group Findings

Tables 1 to 4 present data which define the factors obtained in both contexts when summed across concepts and for the single concept "yourself." Table 1 presents factor loadings and h^2, the scale communalities, (the amount of variance of the scales which can be accounted for by the factors identified) for all concepts in the social studies form of the instrument. Table 2 provides the same data for the literature context. Table 3 and 4 report loadings and communalities of scales on the single concept "yourself" under both conditions of context. These data indicate four major findings: (*a*) the variance accounted for by the factors extracted varies with the context. The greater variance in the social studies context suggests a greater meaningfulness for concepts in this context. Words appear to provide for greater agreement among individuals in the social studies context resulting in higher factor loadings. This may result from the greater objectivity and concreteness of concepts in social studies than in literature. (*b*) The principal dimensions of Evaluation, Potency, and Activity are highly congruent in respect to dimensionality. Although shifts in rank order and differences within the dimensions are indicative of a differential semantic space in each situation. The similarity of principal dimensions in the two contexts suggests that the child considers the meanings of words to be basically the same in each case while the internal differences within the factors indicates that there is a qualitative difference in meaning between contexts. (*c*) The large differences between contexts among the secondary factors, Security, Stability, Tautness, and Novelty also indicate the influence of situational context upon meaning. The secondary factors appear not only to fill in important details as far as connotative meaning is concerned, but may also control, to some degree, the functioning of the principal dimensions. Differences between corresponding factors in the two contexts therefore suggests different responses to the same stimuli. (*d*) Individual concepts show greater variation between contexts than concepts taken together suggesting that summation over concepts produces a regression to the mean. This tends to result in differences between concepts cancelling each other out when concepts are taken together while there may be a marked difference when the concepts are considered individually.

Summary and Conclusions

The theoretical implications of these results are that context "screens" the kinds of situations and symbols which result in meaning for words. Thus words come to have qualitatively different meanings in different situations (contexts) and language can be said to mirror a different kind of reality in each of these different contexts. This conclusion provides support for the linguistic *Weltanschauung* position taken in this study.

The principal dimensions identified can be considered fundamental attributes of an individual's affective functioning. Carroll (1959), in examining dimensions of meaning revealed by the Semantic Differential has proposed that the Evaluative dimension corresponds to a person's tendency to approach or avoid a stimulus. The Activity dimension refers to the necessity or non-necessity for making movements in adjusting to stimuli. Potency, then, suggests the amount of adjustment to be made.

Similarly for the secondary dimensions it appears that the Novelty dimension is an indication of the individual's examination of the concept in terms of its familiarity. Stability becomes that dimension which indicates the changeability of concepts and the necessity of adjusting to changed stimuli. Tautness is suggestive of alertness and tension in regards to response to verbal stimuli. Finally, the Security dimension corresponds to the individual's total "set" and the necessity of response to achieve equilibrium. Here, then, is the value of Semantic Differential indexing of meaning. Factors produced by analysis of SD data indicate the kind of behaviour patterns resulting from the individual's encounter with verbal stimuli. Since it has been shown that meaning patterns vary with the context, it is concluded that behaviour patterns will vary similarly.

While the principal dimensions can be considered fundamental attributes of psychological functioning, the secondary dimensions might be regarded as the catalytic influence bringing about the approach-avoidance response described by the principal dimensions. This postulate would make these dimensions of meaning crucial in the individual's behaviour complex and add a degree of support to the language relativity hypothesis proposed by Whorf. Responses to words, therefore, appear to be in part on the basis of characteristics as identified by their evaluative, potency, and activity dimensions and triggered by psychical components such as novelty and stability.

The nature of the factors obtained and the theoretical implications suggest the presence of underlying constructs or sets which screen incoming symbolic data to generate specific meaning patterns for concepts. These sets are developed through different language experiences resulting from differential situational contexts. Subsequently these same psychological processes "unwind" to produce appropriate responses to verbal stimuli. These sets or mediated response patterns are believed

TABLE 1

ROTATED LOADINGS AND COMMUNALITIES OF THE 20 TEST SCALES*

SUMMATION ACROSS CONCEPTS - SOCIAL STUDIES

Scale	h^2	Factors				
		Evaluation	Activity	Potency	Novelty	Stability
clean - dirty	.565	.666				
painful - pleasurable	.631	.780				
soft - hard	.521	-.439		.541		
light - heavy	.534			.686		
rough - smooth	.540	-.494		.533		
moving - still	.497		.661			
strong - weak	.537		.612			
fast - slow	.545		.716			
important - not important	.450	.514				
kind - cruel	.684	.814				
excitable - calm	.338					-.502
large - small	.368			.536		
dangerous - safe	.576	.692				
bad - good	.694	.812				

steady - changing	.721					-.843
unusual - usual	.635				.778	
ugly - beautiful	.567	.751				
hot - cold	.419		.573			
complex - simple	.319			.412		
uncertain - certain	.539				.623	
Percent of Total Variance	53.401	21.784	9.885	9.145	6.752	5.835

*Loadings of .400 or higher are taken to define the factors obtained.

TABLE 2

ROTATED LOADINGS AND COMMUNALITIES OF THE 20 TEST SCALES*

SUMMATION ACROSS CONCEPTS - LITERATURE

Scale	h^2	Factors				
		Evaluation	Potency	Oriented Activity	Novelty Stability	Stability
clean - dirty	.555	.556	-.440			
painful - pleasurable	.612	.754				
soft - hard	.512		.631			
light - heavy	.527		.715			
rough - smooth	.505		.617			
moving - still	.560			.716		
strong - weak	.549			.669		
fast - slow	.551			.734		
important - not important	.465	.591				
kind - cruel	.660	.761				
excitable - calm	.279				.455	
large - small	.440					-.472
dangerous - safe	.568	.696				

bad - good	.677	.793				
steady - changing	.780					-.819
unusual - usual	.376				.500	
ugly - beautiful	.534	.709				
hot - cold	.302				-.409	
complex - simple	.542				-.629	
uncertain - certain	.483				.579	
Percent of Total Variance	52.382	19.845	9.740	9.318	7.698	5.780

*Loadings of .400 or higher are taken to define the factors obtained.

TABLE 3

ROTATED LOADINGS AND COMMUNALITIES OF THE 20 TEST SCALES*

CONCEPT "YOURSELF" - SOCIAL STUDIES

Scale	h^2	Factors						
		Evaluation	Activity	Potency One	Potency Two	Activity Two	Security	Stability
clean - dirty	.495	.574						
painful - pleasurable	.676	.716						
soft - hard	.712				-.788			
light - heavy	.671			.766				
rough - smooth	.576				-.550		.403	
moving - still	.688					-.793		
strong - weak	.733		-.832					
fast - slow	.664		-.731					
important - not important	.496	.523						
kind - cruel	.578	.625						
excitable - calm	.558						-.698	
large - small	.708			.815				
dangerous - safe	.582	.637						

bad - good	.627	.738						
steady - changing	.666				-.408			.518
unusual - usual	.603	.427						-.401
ugly - beautiful	.462	.541						
hot - cold	.515		-.493			.413		
complex - simple	.682							.781
uncertain - certain	.593						-.536	
Percent of Total Variance	61.427	17.288	10.314	7.374	6.782	6.714	6.668	6.286

*Loadings of .400 or higher are taken to define the factors obtained.

TABLE 4

ROTATED LOADINGS AND COMMUNALITIES OF THE 20 TEST SCALES *

CONCEPT "YOURSELF" - LITERATURE

Scale	h^2	Factors					
		Evaluation	Activity	Novelty	Security	Tautness	Potency
clean - dirty	.595						-.570
painful - pleasurable	.388	-.519					
soft - hard	.541	.572					
light - heavy	.460			.536			
rough - smooth	.632				.778		
moving - still	.573		.502	.452			
strong - weak	.650		-.766				
fast - slow	.667		-.781				
important - not important	.558	-.707					
kind - cruel	.535	-.492					
excitable - calm	.579					-.676	
large - small	.491						.633
dangerous - safe	.614			.469	-.487		
bad - good	.560	-.531					

steady - changing	.562			-.703			
unusual - usual	.424			.473			
ugly - beautiful	.495	-.682					
hot - cold	.640					.658	
complex - simple	.517				.520		
uncertain - certain	.571	-.517		.441			
Percent of Total Variance	55.270	13.521	10.231	9.511	8.266	6.905	6.837

*Loadings of .400 or higher are taken to define the factors obtained.

to represent differential views of reality as a function of language. Language, therefore, becomes the mold into which the individual is cast.

These results have implications for instruction. Since the connotative meaning of concepts appears to vary with the context, suggesting a differential perception of reality, educators must be aware of the set created. If the individual's world view screens or codes environmental stimuli in terms of the meaning symbols held, then conscious effort must be made to either enrich those perceptions which are accurate or present evidence to alter inaccurately structured concepts. Active searching on the part of the learner and an inductive approach to concepts should prove valuable in developing word meanings.

The specific nature of semantic space in different contexts should provide clues as to appropriate instructional materials and procedures. For example, if from a social studies view concepts are more varied in meaning, then materials and procedures designed to present the varied characteristics of these concepts should be employed.

If the secondary dimensions identified in this study serve as catalysts to activate the fundamental attributes of human behaviour represented by the principal dimensions, careful attention must be given to the development of these internal states for without them meaning in its fullness may not be released. This means giving attention to the way the individual feels and thinks about the concept as well as simply trying to define it for him.

Finally, the fact that meaning is more personal in certain contexts such as literature, should cause teachers to plan for the development of concepts along individual lines without emphasizing a particular correctness.

References

Carroll, J. B. Review of Osgood, Suci, and Tannenbaum, The measurement of meaning. *Language,* 1959, 1959, 35, 58-77.

Eastman, M. N. G. A semantic differential analysis of the concept "school." Master's Thesis, The University of Alberta, 1963.

Jenkins, J. J., Russell, W. A., & Suci, G. J. An atlas of semantic profiles for 360 words. *American Journal of Psychology,* 1958, 71, 688-699.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika,* 1958, 23, 187-200.

Lilly, R. S. A developmental study of the semantic differential. Unpublished doctoral dissertation, Princeton University, 1965.

Norman, W. T. Stability characteristics of the semantic differential. *American Journal of Psychology,* 1959, 72, 581-584.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. *The measurement of meaning.* Urbana: University of Illinois Press, 1957.

Saporta, S. (Ed.) *Psycholinguistics, a book of readings.* New York: Holt, Rinehart & Winston, 1961.

Thurstone, L. L. *Multiple-factor analysis.* Chicago: University of Chicago Press, 1947.

Alberta Journal of Educational Research
Vol. XV, No. 1, March, 1969

The authors present three studies of the effects of variations in transformations and in materials on conservation of number in young children. Variations in transformations seemed to have no effect upon performance. Differences in materials appeared to interact with differences in subjects.

DONALD L. PETERS
and
KENNETH RUBIN
The Pennsylvania State University

The Effects of Cued Materials and Transformation Variations and Conservation of Number Performance[1]

In the typical conservation of number task (Piaget, 1952), the subject is questioned regarding the numerical equivalence of two sets of objects before and after a spatial rearrangement (transformation) of one of the sets. The transformation is designed to emphasize the independence of the number dimension of the experimental set, i.e., the number of objects it contains, from the length dimension, the linear space covered by the aggregate. The former is relevant for decisions concerning the equivalence of the two sets and the latter is irrelevant and misleading. The subject must discriminate the appropriate cue dimension if he is to discover and assert the conservation principle that the number of objects in the array remains invariant through transformations of the spatial arrangement.

It has been suggested (Peters, 1967; Zimiles, 1963, 1966) that some aspects of the procedures and materials used in current conservation studies tend to obscure the independence of the length and number dimensions, and to direct the subject's attention to the misleading length dimension. Other procedures and materials may perform the opposite facilitative function. As such, the choice of procedures and materials used may introduce an undesirable, and often unacknowledged, source

[1] A study of the effect of variations in transformations in the context of length conservation was reported by Sawada and Nelson in this journal—Vol. XIV, No. 2, June, 1968. (Ed.)

of variance into the testing situation which makes interpretation across studies and inferences to new situations difficult.

Problem

This report presents a series of three studies dealing with conservation of number. The hypothesis investigated is the null hypothesis that (*a*) variations in the transformations used and (*b*) the variations in the materials will have no effect on conservation of number performance. Although alternative verbal testing procedures also contribute to the variance, these have been studied and discussed elsewhere (Pratoomraj and Johnson, 1966; Mermelstein and Shulman, 1967) and were not included as variables in this series of studies.

Methods

Sample

A total of 93 lower socio-economic preschool and kindergarten children were used in the three studies. The subjects for Studies One and Two were randomly selected from three kindergarten classrooms in the San Francisco Bay area of California. Study Three's subjects were obtained from a preschool Head Start program in Trenton, New Jersey. Relevant information concerning the sample for each study is found in Table 1.

TABLE 1

CHARACTERISTICS OF THE SAMPLES

Study	Sex	Mean Age	SD	Caucasian	Negro	Other
1	Male	68.0	3.54	11	2	2
	Female	68.1	4.12	9	1	5
	Total	68.0	3.84	20	3	7 /30
2	Male	67.6	4.34	10	3	5
	Female	67.6	3.26	9	4	2
	Total	67.6	3.89	19	7	7 /33
3	Male	61.8	3.73	9	5	1
	Female	59.5	2.75	13	1	1
	Total	60.7	3.47	22	6	2 /30

Measures

The conservation of number task utilized an 18" x 30" testing board divided by a black line to provide two 8¾" x 30" testing fields, one red and the other blue. Four sets of testing blocks were used at different times (See Table 2). Material Set 1 was considered neutral in that the materials provided no cues highlighting either the relevant or irrelevant dimensions. Material Set 2 was thought to provide cues for correspondence matching and also to elicit a counting response. Material Set 3 provided only cues for correspondence matching. Material Set 4 provided no correspondence matching cues but, because of the dimensions of the blocks, was thought to accentuate the length dimension. The child was requested to "Make a bunch of blocks on the red side that has the same number of blocks as the bunch on the blue side."[2]

TABLE 2

DESCRIPTION OF MATERIALS*

Block Set #	Set Names	Dimensions	Color
1	Neutral	1 3/4" X 1 3/4" X 7/8"	Yellow.
2	Number Accentuated	1 3/4" X 1 3/4" X 7/8"	White blocks with black sequins ordered 1 through 9.
3	Correspondence Accentuated	1 3/4" X 1 3/4" X 7/8"	Red, blue, yellow, dark green, light green, white, silver.
4	Length Accentuated	3 1/2" X 1 3/4" X 7/8"	White block with five evenly spaced black sequins.

* each set included 18 blocks.

The initial correspondence sequence was followed by a transformation of the display on the experimenter's (blue) side and a standard questioning procedure. The board was cleared and the sequence was repeated with a new transformation introduced on the subject's side of the display. The standard procedures are presented in detail elsewhere (Peters, 1968). The aggregate size was six in all cases. The following rating scale was used to evaluate the child's performance:

1. The child did not understand the problem. He used all blocks without regard for the formation to be matched.

2 The authors wish to acknowledge the major contribution of Dr. Barbara Rothenberg of the Educational Testing Service to the design and standardization of the testing and questioning procedures used in these studies.

2. The child could not make the correct one-to-one correspondence on either occasion within the trial. He continually focused upon the length or amount of space covered by the model.
3. The child was inconsistent within a trial on making the initial one-to-one correspondence, or he was illogical in his sequence of responses concerning the initial equivalence of the two aggregates. For example, in response to, "Does this bunch on the red side have the same number of blocks as this bunch on the blue side?" and, "Does one bunch have more blocks?" he responds, "Yes," "Yes."
4. The child established initial one-to-one correspondence and was consistent in his responses concerning initial equivalence, but he became confused as the visual display was transformed. He gave non-conserving responses concerning the equivalence of the two aggregates after transformation.
5. The child established initial equivalence, but was inconsistent (sometimes conserving, sometimes not) on questions concerning the equivalence of the two aggregates after transformation.
6. The child established initial equivalence through one-to-one correspondence and continued to assert the equivalence throughout the transformations. However, he could not formulate an adequate verbalization of the rule for conservation as justification for his assertions.
7. The child displayed full conservation and adequately verbalized a rule for conservation. To be considered an "adequate" verbalization, the child's response had to include a statement of equivalence based upon (*a*) counting, (*b*) the lack of addition or subtraction of new members, or (*c*) the reversibility of the transformation.

Procedures specific to the individual studies are provided in the description of each.

Study One

In order to determine the effects of introducing alternative transformations and cued materials, it was necessary to obtain reliability estimates for the effects of scoring and retesting. The latter served both as a measure of the stability of the behaviour and as a baseline record from which to evaluate improvement.

The analysis employed was designed to determine the magnitude of the discrepancies likely to arise from multiple facets of the measuring procedures. The generalizability model (Cronbach, Rajaratnam, and Gleser, 1963; Gleser, Cronbach, and Rajaratnam, 1965) suggests that every measuring operation is a sample from a universe of admissible operations. The resulting measurement departs to some degree from the universe score—the average across all admissible observations. The generalizability model provides a means of estimating the magnitude of the discrepancies likely to arise under a given measurement procedure, and provides formulae for establishing interval and point estimates of the universe score. Further, it is possible to obtain simultaneously independent estimates of the contributions of multiple facets of the

measuring operation, each of which contributes to the discrepancy between the observed score and the universe score.

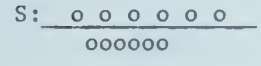
Thirty subjects (p) were administered the conservation of number measure on three occasions (o) using Material Set 1. The transformations employed are indicated by asterisks in Table 3. For each occasion the protocols were evaluated by two independent scorers (j). This led to a completely crossed design (p x j x o). Variance components for each source of variance were computed separately by sex and the coefficient of generalizability ($E\rho^2$) was estimated by dividing the estimated universe score variance by the expected observed variance. The results are presented in Table 4.

TABLE 3

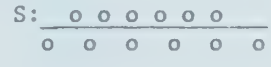
CONSERVATION OF NUMBER TRANSFORMATIONS

Linear Transformations

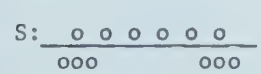
1.* Row Contraction



2.* Row Expansion

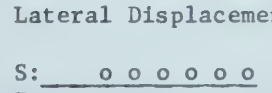


3.* Equal Subdivision

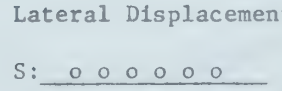


4.* Unequal Subdivision

5. Lateral Displacement

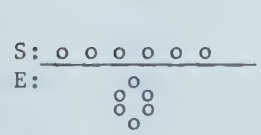


6. Lateral Displacement



Horizontal Figure Formation

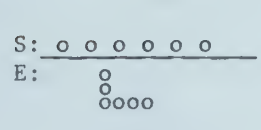
7.



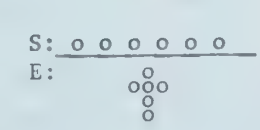
8.

9.*

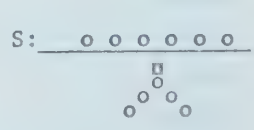
10.



11.

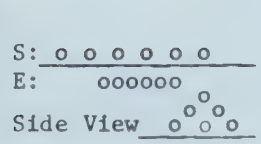


12.

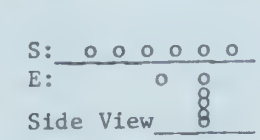


Vertical Plane Figure Formation

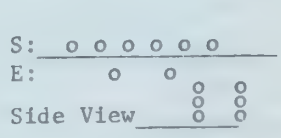
13.*



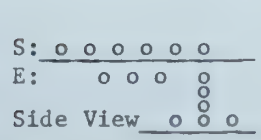
14.



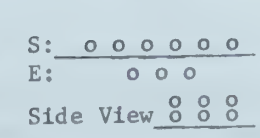
15.



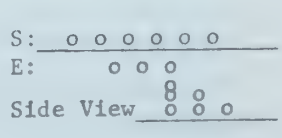
16.



17.



18.



$E\rho^2$ is an intraclass correlation which has been variously defined as (*a*) the squared correlation of the universe score with the observed score, (*b*) the expected intraclass correlation between pairs of observations by repeating the design more than once, and (*c*) the percentage of observed variance that is, on the average, accounted for by the universe score (Cronbach, 1967). $E\rho$ is, therefore, an approximation of the correlation to be expected if the study were repeated with different scorers and on different occasions.

TABLE 4

STUDY ONE MEANS, STANDARD DEVIATIONS, AND VARIANCE COMPONENTS

	Occasion 1		Occasion 2		Occasion 3	
	Scorer 1	Scorer 2	Scorer 1	Scorer 2	Scorer 1	Scorer 2
Males	3.93	3.93	4.26	4.26	4.93	4.86
(N = 15)	0.68	0.68	0.77	0.77	1.38	1.31
Females	4.06	4.00	4.46	4.53	4.53	4.53
(N = 15)	0.85	0.89	1.35	1.41	1.41	1.41

Source of Variance	df	Mean Squares	Component of Variance	Freq. within p	Cont$_2$ to $E\sigma^2$ (xp)	$E\rho^2$	$E\rho$
Males: N = 15							
Subjects (p)	14	3.2905	0.4288	1	.4288	.48	.69
Scorers (j)	1	0.0111	0.0000	1	.0000		
Occasions (o)[a]	2	7.2333	0.1936	2	.0968		
jp	14	0.0111	0.0000	1	.0000		
oj	2	0.0111	0.0000	2	.0000		
op	28	1.4238	0.7064	2	.3532		
Residual (ojp, e)	28	0.0111	0.0111	2	.0055		
Females: N = 15							
Subjects (p)	14	5.7825	.8365	1	.8365	.64	.80
Scorers (j)	1	0.4000	.0056	1	.0056		
Occasions (o)[a]	2	3.1444	.0952	2	.0476		
jp	14	0.7095	-.0278	1	b		
oj	2	0.2333	-.0373	2	b		
op	28	0.8468	.0270	2	.0135		
ojp, e	28	0.7928	.7928	2	.3964		

The results attest to the stability of performance and the objectivity with which the scoring may be accomplished. The contribution to the variance made by two scorers of highly dissimilar levels of training was

essentially zero. The main effect across occasions was significant at at least the .05 level for both sexes, but the ordering of subjects was reasonably stable.

Study Two

To determine the effects of variations in the transformations used in evaluating number conservation performance, 33 kindergarten children were tested twice using linear transformations, twice using horizontal figure transformations, and twice using vertical figure transformations. The specific transformations confronting each subject were randomly assigned from those pictured in Table 3, with two separate transformations used in each testing. In all cases the order of testing was linear, horizontal and vertical. All testing was done with Material Set 1. The average scale score for the two testings within each type of transformation was used as the subjects' score in the analysis.

The results (see Table 5) indicate there is a considerable similarity between the ordering of subjects in testings using any of the three types of transformations. Approximately 30% of the variance in performance is attributable to the combination of the transformations, the repeated testing, and residual error. Table 4 indicates retesting and its interactions account for approximately 14% of the variance. It would seem, therefore, that the use of different transformations makes about the same contribution to the variance.

TABLE 5

CORRELATION MATRIX FOR STUDY TWO

Transformation Type	Mean	SD	Linear	Horizontal	Vertical
Linear	4.77	1.33	1.00	0.84	0.80
Horizontal	4.73	1.42		1.00	0.86
Vertical	5.15	1.48			1.00

The mean score for the vertical transformations is significantly higher ($p<.05$) than those for the other transformations. However, again using the information provided by Study One, the significant increase may be accounted for by the learning or test-wiseness resulting from repeated testing.

Study Three

Thirty preschool children were pretested using transformations one and two and Material Set 1. They were randomly assigned to five groups. All groups were retested the following day with the same transformations, but with different sets of materials. Group 1 was retested

with Material Set 1 (neutral materials), Group 2 with Material Set 2 (number accentuated), Group 3 with Material Set 3 (correspondence accentuated), Group 4 with Material Set 4 (length accentuated), and Group 5 with Material Set 3. Groups 3 and 5 were combined for the analysis.

The results are presented in Table 6. All groups showed significant improvement between the two testing, ($p<.01$, $\chi^2 = 11.3$). No significant differences were found among the mean scores of the groups.

TABLE 6

STUDY THREE RESULTS

		Occasion 1		Occasion 2		Coefficient
Group	Materials	Mean	SD	Mean	SD	of Stability
1	Neutral	2.16	0.68	2.66	0.74	+0.75
2	Number Accentuated	2.16	0.37	2.66	0.47	+0.31
3 + 5	Correspondence Accentuated	2.08	0.49	2.58	0.75	+0.09
4	Length Accentuated	2.00	0.57	2.50	0.50	+0.58

The correlation between the two testings, separately by groups using different materials, is also presented. Where the materials were uncued, and remained unchanged between testings, results highly similar to those found in the equivalent testing situation of Study One were found. Where the materials were varied between the two testings, a decrease in the correlation is evident. A comparison of the correlations for Groups 3 and 5 combined and Group 1 approaches significance ($p<.08$). The between testings correlations for Groups 2 and 4 are in an intermediate range. The stability of performance of individual subjects seems to vary considerably depending upon the materials used in the testing.

Discussion

For each of the studies, a standard set of procedures was used. The wording of questions, and the sequence of events in the testing situation were identical. In each, the presentation of reversible transformations led to an increase in score across the repeated testing dimension. When the transformations and the materials used in the sequence were held constant, the performance of the subjects on subsequent testings remained stable yielding correlations in the .69 to .86 range. Such coefficients of stability are highly similar to those commonly found in

the literature (Dodwell, 1960; Wallach and Sprott, 1964, Gruen, 1966; Peters, 1967).

Introducing variations in the transformations used in the testing seems to have no effect upon the subject's performance. The data offer no justification for the rejection of the null hypothesis. Test results using linear, horizontal, and vertical transformations were found to be highly correlated indicating a stability in performance across transformations.

Some support has been found for the rejection of the hypothesis that variations in the materials would have no effect on conservation of number performance. Since the procedures and transformations in testing were controlled in Study Three, the variation in the correlations found for Groups 2, 4, and 3 and 5, may be attributed to the use of different materials for the two testings.

It is clear from the data presented in Table 6 that the differential effect of the materials is not manifested in a mean difference of scores for the groups. It appears instead that differences in materials interact with individual differences in subjects. The low correlation ($r = .09$) found for Groups 3 and 5 combined reflect a change in the rank order of individuals from one test trial to the next. Presumably, variations in the cues provided by the materials were attended to by some subjects and not by others. Further, for some subjects the cues may have served to facilitate appropriate behaviour, for others they may have served as distractors.

The findings of Study Three suggest that some care should be taken as to the kinds of inferences that can be made across studies. Of particular importance is the possibility of gaining additional insight into the processes involved in transition between non-conservation and conservation by means of investigating in detail the use of cued materials of various kinds and of the individual differences in children in response to these materials.

References

Cronbach, L. J. Personal communication, 1967.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. Theory of generalizability: A liberation of reliability theory. *British Journal of Statistical Psychology,* 1963, 16, 137-163.

Dodwell, P. C. Children's understanding of number and related concepts. *Canadian Journal of Psychology,* 1960, 14, 191-205.

Gleser, G. C., Cronbach, L. J., & Rajaratnam, N. Generalizability of scores influenced by multiple sources of variance. *Psychometrika,* 1965, 30, 395-418.

Gruen, G. Experience affecting the development of number conservation in children. *Child Development,* 1965, 36 (4), 963-979.

Mermelstein, E., & Shulman, L. Lack of formal schooling and the acquisition of conservation. *Child Development,* 1967, 38 (1), 39-53.

Peters, D. L. Task variations and individual differences in Piaget's conservation of number. *Merrill-Palmer Quarterly,* 1967, 13 (4) 243-258.

Peters, D. L. Piaget's conservation of number: The interaction of language comprehension and analytic style with three methods of training. Unpublished doctoral dissertation, Stanford University, 1968.

Piaget, J. *The child's conception of number.* New York: Humanities Press, 1952.

Pratoomraj, S., & Johnson, R. Kinds of questions and types of conservation tasks as related to children's conservation responses. *Child Development,* 1966, 37 (2), 343-355.

Wallach, L., & Sprott, R. Inducing number conservation in children. *Child Development,* 1964, 35, 1057-1071.

Zimiles, H. A note on Piaget's conception of conservation. *Child Development,* 1963, 34, 691-695.

Zimiles, H. The development of conservation and differentiation of number. *Child Development,* 1966, 31 (Whole No. 6).

Alberta Journal of Educational Research
Vol XV, No. 1, March, 1969

The author describes the ecological aspects of education in a small Canadian city and discusses the usefulness of such information in educational decision-making.

HELEN L. PIERCE
The University of Alberta

Ecology of Education

The ecology of education is an increasingly important concern of educators, civic officials and professional planners. Two major workshops on the ecology of the classroom have been conducted recently by the National Education Association.[1] In urban design the optimum size and location of schools are major problems in rapidly growing cities (1968). Leading thinkers in education with their ideas of the mobile student, the mobile family, the mobile school, the educational park, the Great High Schools and the multi-campus high school, reflect increasing concern with spatial or ecological aspects of education (1967, 1968). This article discusses the nature of ecology as a basic social process and presents a brief report on ecological aspects of education in a small Canadian city, Camrose, Alberta (1964).

Ecology Explained

Ecology is the study of the distribution in space of living organisms—plants, animals, and humans—and is concerned with the processes by which adaptations are reached and maintained. These ecological processes (concentration, centralization, sub-centralization, segregation, invasion, and succession) are specialized forms of the competitive process.

Human ecology has to do with the patterning of regions, communities, and other groupings. It explains the distribution of human beings and their institutions in space at a given time and, in the duration of time, it provides systematic knowledge of social change. It includes symbiotic dependence, inter-dependence resulting from communication and mutual understanding, and diffusion of knowledge. Competitive selection takes place, locations are subject to the main trends of growth and change, institutions emerge where there is need for their services, and they compete spatially with others. Distribution is not consciously planned

[1] Research Council of the Association for Supervision and Curriculum Development, NEA, Washington, D.C. sponsored research institutes on "The Ecology of the Classroom" at Washington, D.C., November, 1968 and San Francisco, April, 1969.

in many situations. Usually the planning, if any, comes after problems of competition for space become evident.

Human ecology, as a field of study, has been developed primarily by sociologists who have been theoretically oriented. However, the applications of ecology are especially significant for those most closely associated with the spatial development of communities—town planners, administrative bodies, officers of community institutions, and professional educators. Though the situation is now changing, the ecology of education has been an underdeveloped area of knowledge.

Camrose—City of the Study

Camrose, a trade city, is located in Central Alberta parkland in a humid, continental cool summer climate with extremes in temperatures. The economy of the area is based largely on agriculture with some coal, oil and small industries. There is access to the city, located about forty-five miles southeast of Edmonton, by rail and by highways from all directions. Camrose received city status in 1955 after growing from a hamlet to a city in fifty years, and by 1961, when this research was begun, the population had reached 6,924. While some town planning was started when the urban centre was still a town, the Battle River Planning Commission became an advisor when Camrose received city status.

The school facilities grew from accommodation in a temporary classroom to multi-room and multi-system. The Camrose Public School District, the Camrose Separate School District, and the Camrose School Division operated schools in the city. There were, in addition, three private schools under the auspices of churches. The provincial normal school, opened in 1912, was closed in 1938. Proposed building programs in 1962 indicated further expansion of the school facilities.

Findings

The findings of the study are in three categories. First, the location, acquisition, and area of school sites are considered. Second, the distance journeyed to school, the mode of journey to school, and the noon lunch situation are examined. Finally, the ecology of the school as a factor in school and community relations is summarized.

The location, acquisition, and area of school sites seem significant to both town and school administration. It was found that school sites tended to be on the periphery when they were occupied first. Only two—those purchased in 1957 and 1962—of the fifteen sites were not acquired from the town. Moreover, until 1957 school sites were provided free or nearly free by the town. Contrary to what was hypothesized, there was no regular or consistent increase in areas of sites per regular classroom in schools. Instead, the enlargement of three school buildings decreased

the site area per regular classroom from what it had been originally. Those sites acquired in 1962 showed an increase in size over sites acquired previous to that time. However, among the private schools, the one established first had the largest site while the one established last had the smallest site. Briefly, the trend seemed to be toward purchased and larger peripheral sites with only the school sites selected in 1962 reaching the recommended area of the Battle River Planning Commission.

While several recommended school site sizes had been made for Alberta schools, the Battle River Planning Commission had a simple guide. It recommended that the school site should have one acre and not less than two-thirds of an acre per classroom as a minimum. Also, it suggested that in urban centres one elementary school per section of land should be considered when planning land use in cities. However, other writers concerned with school-site size recommended less than one acre per classroom for the school grounds.

Simple schedules were distributed to and completed by all students in the school systems studied to determine the address, mode of journey to and from school, and the number of times daily that the journey was made for each student in the months of October, January, and May. While the findings indicated that the distance travelled varied by grade level and school systems, it was found that four-fifths of all school personnel lived within three-quarters mile of school. There were sixty-five percent of the elementary school pupils living less than one-half mile and eighty-seven per cent living less than three-quarters mile from school. Only one school had all its pupils living within three-quarters mile. At the junior high school level, there were seventy-eight per cent of the pupils living within three-quarters mile of the school attended and eighty-nine per cent living less than one mile from school. Senior high school students travelled slightly further than junior high school students—sixty-eight per cent of the students lived within three-quarters mile and eighty-one per cent lived within one mile of the school. The remainder of the senior high school students lived beyond the one-mile distance limit in the city or came from the rural area.

In general, distances travelled to school by junior and senior high school pupil personnel compared favorably with recommended maximum distances suggested in literature, but for elementary schools some pupils travelled further than suitable maximum distances. Also, the problem of distance was greater for the Separate School District, with one elementary school, than in the Public School District with six elementary schools. Recommended maximum distances varied for elementary school pupils from one-quarter to three-quarters mile, and for junior high school pupils from one to one and one-half miles, while for senior high school students the recommended distance was two miles.

While one might expect that traffic arteries and natural boundaries

would determine school attendance boundaries, it was found that in Camrose this was not always the case. Rather, it seemed that traffic arteries and natural boundaries influenced but did not define attendance boundaries.

The journey to school can present a problem when one considers the distance factor and the climatic conditions of Alberta. In Camrose, the family assumed the major responsibility for children getting to and from school. Generally, most pupils journeyed to and from school independently, with dependence on parents and others varying with the season and time of day. Dependence upon others for conveyance was least in the fall and spring and in the afternoon journey from school.

While data were not analysed so as to test such a hypothesis, findings suggested that the mode of journey to school was correlated with the distance to school that the pupils travelled and with the season. Maximum distances to travel independently to school seemed apparent at different grade levels and for different seasons. For elementary school pupils, it seemed to be from one-half to three-quarters mile in October, from one-quarter to one-half mile in January, and about three-quarters mile in May. For junior high school pupils the maximum distance travelled independently in October seemed to be three-quarters mile to school and three-quarters to one mile from school, in January one-half to three-quarters mile to school and one-half to one mile from school, and in May three-quarters to one mile. For senior high school pupils the maximum distance travelled independently appeared to be one and one-quarter miles in October and May and three-quarter to one mile in January.

Then, if pupils must travel to school independent of others, is the maximum distance for them to travel the maximum for the month of January—for elementary school pupils from one-quarter to one-half mile, for junior high school pupils one-half to three-quarters mile, and for senior high school pupils three-quarters to one mile? It is noted that these distances are less than those recommended as grade-level distance maximums to which reference was made earlier.

The number of pupils having noon lunch at school seemed to be influenced by several variables. Findings suggested that elementary school pupils living beyond three-quarters mile from school in October, beyond one-half mile in January, and beyond one mile in May ate their noon lunch at school. At the junior high school level, it appeared that pupils living more than one mile from school in October, more than three-quarters mile in January, and in the rural area only in May ate their lunch at school. The maximum distance for senior high school pupils to travel home for noon lunch seemed to be approximately one mile in October and May and slightly less in January. In addition, variations by schools of the same grade level suggested that socio-economic factors

may have influenced pupils to eat lunch at school—that parents were away from home during the noon recess.

By means of interviews with school principals, the use by the public of school facilities was determined. Facilities available and the church and/or school associations as well as the location of the school appeared to affect the use made of schools by the adult community. Schools used for meetings were located both on or near the periphery and in the central part of the city. Each school system, with the exception of the Public School District which had two schools used, had one school which was used as a meeting place for adults. Also, those groups associated with the school through their children or their church tended to make the most use of the school property.

Some school facilities in the city were used for adult activities other than meetings. All students enrolled in adult education offered in schools in Camrose were asked to complete a simple schedule. It was noted that while nineteen per cent of those enrolled stated that the nearness of the school was a factor in deciding to take the course, the analysis did not indicate whether the location of the two schools where courses were offered was a factor in other adults not enrolling in a course. For concerts and sports, the space and facilities available rather than the location of the schools seemed to be the factors determining the demand for the use of schools. For certain miscellaneous uses of schools, the space, facilities, and church affiliation of the sponsoring group appeared to determine the requests for use of school property.

To summarize, school sites, provided free or very nearly free by the town until 1957, were located on the periphery. The trend seemed to be toward purchased and larger peripheral sites. There were four-fifths of all school personnel living within three-quarters mile of school. Recommended maximum distances were exceeded only by some elementary school pupils. While most pupils journeyed to school independently, the maximum distance of independent journey appeared to be from one-quarter to three-quarters mile for elementary, one-half to one mile for junior high, and three-quarters to one mile for senior high school pupils. The analysis suggested that the critical distance associated with eating noon lunch at school is one-half mile for elementary school pupils and one mile for junior and senior high school pupils. School location and facilities seemed to influence some but not all community use of schools.

Discussion

The summarized findings seem to have certain implications for school and community administration. Some that appear evident to the writer are listed below. Others may have been suggested to the reader.

1. A consideration of the ecology of education in a town or small

city by educational planners can contribute to the achievement of the goals of education. Moreover, while town and educational planners or administrators may find particular concern with specific aspects of the ecology of the urban centre, there is much which needs the attention of both bodies of administrators. It seems essential that these administrative personnel be familiar with the ecology of school sites of school personnel, and of schools as a factor in school and community relations.

2. City and school administrative bodies, as co-operating agencies, if they are familiar with the ecological processes and observe them at work in their city, can use these observations to assist in city and educational planning. For example, it seems that ecological processes of invasion and succession should be investigated before educational facilities are expanded or renovated, that educational planners should be familiar with projected plans of the city growth since future school population location will be influenced, and that educational planners should avoid construction on already occupied school sites which will reduce the area of the playground to a size too small for the school enrolment. The Battle River Planning Commission, an advisor to the City of Camrose, includes plans for new school sites in residential development areas which seem worthy of serious consideration by educational planners. Also, if traffic arteries and parks are located in relation to natural boundaries and divide the city into areas large enough to require at least a six-room public elementary school, and if development of long narrow residential areas are avoided, then attendance areas for schools can be established without hazards to children from heavy-traffic thoroughfares.

3. Study of the ecology of school personnel and observation of the ecological processes by city and school administrative bodies can assist in the administrative tasks. If administrative bodies consider the probable distance school pupil personnel journey independently to school, it should assist in selecting new school sites and determining attendance boundaries. Where city transit service is not available, administrative bodies could aim to locate schools to enable pupils from less desirable residential areas to be within the optimum distance to travel to school independently. (From these residential areas, families may be unable to provide vehicular transportation for their children, and then distance could be a contributing factor in irregular attendance and withdrawal from school.) School systems may find it necessary to provide accommodation and noon supervision for pupils eating noon lunch at school. Demand for these services are likely to be made if elementary school pupils live more than one-half mile from school, and if junior and senior high school pupils live more than one mile from school. The effect of these maximum distances will vary with school systems, residential areas and seasons.

4. The study of the ecology of school personnel and observation of the ecological processes by school professional personnel could be used to facilitate teaching and learning. Their consideration when dealing with discipline problems, tardiness, school dismissal, curriculum planning, and scheduling of extra-curricular activities could provide a guide in decision making. Ecology of education could shed light on formal and informal organization of pupils, probable influences of social segregation by pupils eating lunch at school, pupils journeying to school by different means, and by rural, urban, or campus residence. It could provide insight into probable influences of invasion on attendance areas by different school systems or by schools within the same school system, and into expenditure of pupil energy in the journey to school.

5. Since the schools are used by the adult community, town and educational planners could consider, from the standpoint of the adult population, the ecology of schools and the ecological processes when planning school sites and buildings and when planning the use of existing school facilities. If school buildings and sites are to be used for such events as concerts, civil defence purposes, and sports activities, city and school administrative bodies could incorporate in school building and site plans suitable auditorium and gymnasium space, shelter areas, and playing fields. Adequate parking space, spectator space, and free access to and from the location then more likely could be assured. For maintenance of shared property, there could be some arrangement for shared responsibility.

6. Cooperative planning seems essential if the ecological aspects of education in our communities are to contribute to the solution rather than the generation of problems. Even though ecology is fundamentally a competitive process, the process can be controlled by cultural means. The most obvious means is through cooperative planning by school and city administrative bodies, planners, and school and city officials.

References

Eldredge, H. W. (Ed.), *Taming megalopolis.* New York: Frederick A. Praeger, 1967.

Havighurst, R. J. Metropolitanism and the schools. *The Urban Review,* 1968, Vol. 3, No. 2.

McDonald, F. J. Beyond the schoolhouse. *The Urban Review,* 1968, Vol. 3, No. 2.

Pierce, H. L. An ecological analysis of the schools in a small western Canadian city: Camrose, Alberta. Unpublished master of education thesis, The University of Alberta, Edmonton, Alberta, 1964.

Robinson, N. W. Marland's "Magnificant Gamble"—Pittsburgh's great high schools. *The Urban Review,* 1968, Vol. 3, No. 2.

Toffler, A. (Ed.), *The schoolhouse in the city.* New York: Frederick A. Praeger, 1968.

BOOK REVIEW

READING TESTS AND REVIEWS

Edited by O. K. Buros. Highland Park, New Jersey: Gryphon Press, 1968.

This is a very welcome compilation of reviews of reading tests taken from the six volumes of the *Mental Measurement Yearbooks,* but also includes reviews of tests published since 1965 when the latest Yearbook was issued. The index includes a comprehensive bibliography of reading tests known to be in print in English speaking countries as of May 1st, 1968, and in addition includes all those now out of print but which were once listed in one or more of the *Mental Measurement Yearbooks.* The pertinent information and appropriate description with respect to each test precede the reviews by reading specialists. References to test reviews appearing in professional journals are also given. For the convenience of the user a classified index to *Mental Measurements Yearbooks,* a directory of publishers and indexes of both titles and names are also included in this volume.

A total of two hundred and ninety two tests are examined. Of these, two hundred and nine are in print and eighty three out of print. The reading tests are classified into eight sections and an analysis reveals the relative numbers in each category: general, 39.7; study skills, 16.4; readiness, 11.6; diagnostic, 11.0; special fields, 7.9; miscellaneous, 5.8; oral reading, 5.5; and speed, 2.1. One interesting sidelight is that although between the closing dates of *The Sixth Mental Measurements Yearbook* and *Reading Tests and Reviews* is slightly less than four years, 42 percent of the two hundred and nine tests currently in print are classified as new, revised or supplemented since the *Sixth Yearbook.* This is an indication not only of the continuing indefatigable activity of reading test constructors, but also of the incessant attempts to include enlarged concepts of the constituents of reading into the assessment of reading behaviours. It may only indicate, however, that reading test publication is continuing to be a profitable commercial investment.

Over the years, "Buros" has become the invaluable vademecum of educational researchers, who have relied upon the *Yearbooks* not only to give the basic test information but also critical assessment of the validity, reliability, limitations and general worth of each test. Undoubtedly these publications have had a more profound effect than that of performing a merely utilitarian and time saving function of easy access to information. Through these reviews, test users have gained greater awareness of values, problems and dangers in the use of standardized tests; test authors and publishers have provided detailed information on the validity and limitations and means of standardization

of the tests; psychometricians and statisticians have been goaded to improve techniques; and the field as a whole has continually reassessed the basic assumptions, beliefs and values which are inherent and relevant to testing. Though it is impossible to document, this incidental contribution to the improvement of techniques and to the evolution of the field of measurement in general, is incalculable.

In the preface, the editor indicates that the plans for this volume were laid twenty-seven years ago, but owing to financial difficulties, publication was delayed. Over a quarter of a century later, the present volume not only fulfils the primary objective of making material more accessible to individuals according to their specific interests, but because of the time span, an historical element is introduced. As a result of this extended period, it is possible to examine the evolution of testing in general and reading in particular. To a student of reading, the shifting, evolving concepts of the field can be clearly demonstrated, for many of the innovations and revisions in reading tests occur as a result of new insights. Though a reading test is in effect the product, underlying assumptions concerning the process may be inferred from the reading behaviours measured. It is significant that over the years, the number and variety of such behaviours measured has increased.

The inclusion of the out of print tests is also of great value to the reading researcher, since it enables him to evaluate more effectively past research based on these tests. With the inclusion of the reviews of revised versions of tests the student of reading can assess the means by which measurement of reading has been sharpened to increase both validity and reliability. Moreover, gaps in the aspects of reading measured are also apparent in this compilation over a period of time. One of the most obvious is that of the paucity of tests which measure critical reading.

Each section is reprinted directly from the original volumes, and consequently does not follow the normal numerical sequencing in pagination. Though it is recognized that this may have resulted from economic stringency, it does detract somewhat from the immediate usefulness of the volume. The specific method of indexing for this volume is not made explicit or given prominence, and consequently the location of particular items is frequently erratic. Yet *Reading Tests and Reviews* is to be welcomed as a useful compendium for workers in the reading field.

Marion D. Jenkinson
The University of Alberta

NOTES TO CONTRIBUTORS

ROBERT D. ARMSTRONG holds bachelor's and master's degrees from The University of Alberta and an Ed.D. from the University of California at Berkeley. He was formerly a teacher and principal in elementary schools of Alberta and is presently an Associate Professor in Elementary Education at The University of Alberta. He is the senior editor of a monograph on Teacher Education, a report of the invitational conference sponsored by the Department of Elementary Education. He has also had articles published in the *Alberta Journal of Educational Research.*

L. W. DOWNEY, before assuming his position as Director of the Alberta Human Resources Research Council, was Head of the Department of Secondary Education at The University of Alberta and Chairman of the Centre for the Study of Administration at the University of British Columbia. He has edited several books and monographs and contributed frequently to journals. His most recent article, "Canadian Image of Education," appears in *Canadian Society.*

PETER EVANECHKO is currently a Ph.D. student in the Department of Elementary Education at The University of Alberta. Prior to coming to The University of Alberta to take a graduate program in language arts he was Supervisor of Elementary Instruction for the County of Leduc. He has published in *One World* the Journal of the Social Studies Council A.T.A.

MARION D. JENKINSON, who received her Ph.D. from the University of Chicago, is a Professor of Elementary Education at The University of Alberta. Prior to her return to The University of Alberta she was Head of the Division of Curriculum Research and Instructional Techniques at the Ontario Institute for Studies in Education. Dr. Jenkinson is well-known for her participation in international conferences of the *International Reading Association* and for her many publications which include journal articles and textbooks.

THOMAS O. MAGUIRE holds B.Sc., B.Ed., and M.Ed. degrees from The University of Alberta and a Ph.D. from the University of Illinois. He is an Assistant Professor of Educational Psychology at The University of Alberta with a joint appointment in the Division of Educational Research Services. His publications include articles on research methodology and curriculum.

DONALD L. PETERS holds a Master of Education degree from Northeastern University and a Ph.D. in Educational Psychology from Stanford University. He is presently an Assistant Professor of Human Development at The Pennsylvania State University. He has served as consultant to the Human Resources Commission of San Mateo County, California and to the San Jose, California Supplementary Education Center. His previous publications have appeared in *Merrill-Palmer Quarterly, Childhood Education,* and the *California Journal for Instructional Improvement.*

HELEN L. PIERCE is an Assistant Professor of Educational Foundations at The University of Alberta where she received her B.Ed. and M.Ed. degrees. Since graduation she has continued study as a special graduate student. Before joining the university staff, she had experience as a teacher in elementary and junior high school grades and as a part-time instructor at the Camrose Luthern College.

KENNETH RUBIN graduated with a B.A. in psychology from McGill University. He is presently a research assistant pursuing a graduate degree in the department of Child Development and Family Relationships at the Pennsylvania State University.

Through an arrangement with New York University Press, publications of The National Foundation for Educational Research in England and Wales will be available in Canada from Copp Clark.

INCLUDED ARE:

THE i.t.a. SYMPOSIUM

Research Report on the British Experiment

by JOHN DOWNING

This is the first definitive report on the famous initial teaching alphabet (i.t.a.). Based on a study of 4,000 children, half using i.t.a. and half using traditional orthography, it offers careful comparisons of methods, results, and conclusions. Included are a number of short papers by prominent reading specialists.

$4.50

STANDARDS AND PROGRESS IN READING

by JOYCE M. MORRIS

This longitudinal report examines the first comprehensive British research into factors influencing children's achievements in reading. These factors include the child himself, the home, and the school, and assesses the attitudes and methods of teachers. Case studies are included.

$7.50

CONCEPT GROWTH AND THE EDUCATION OF THE CHILD

by J. G. WALLACE

A comprehensive survey on the present state of our knowledge in the study of the growth and development of concepts. The first section deals with research in the traditional behaviour theory and the second is mainly devoted to the work of Piaget and his associates.

$6.00

THE COPP CLARK PUBLISHING COMPANY

517 Wellington Street West
TORONTO 2B, ONTARIO